

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 01:50:51 ; Search time 3558.48 seconds

11163.571 million cell updates/sec

Title: US-09-986-682B-4

Sequence: 1 ATGACACAGCGGGACTACAA.....CCGGACAAGACAA TTCCA 1365

Scoring table: IDENTITY_NUC

Gapor 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

[illegible]

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1:	gb_ba:*
2:	gb_hlg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vt:*
15:	en_ba:*
16:	en_fun:*
17:	en_hum:*
18:	en_in:*
19:	en_mn:*
20:	en_om:*
21:	en_or:*
22:	en_ov:*
23:	en_pat:*
24:	en_ph:*
25:	en_pl:*
26:	en_ro:*
27:	en_sts:*
28:	en_un:*
29:	en_vt:*
30:	en_hlg_hum:*
31:	en_hlg_in:*
32:	en_hlg_om:*
33:	en_hlg_ov:*
34:	en_hlg_pat:*
35:	en_hlg_ph:*
36:	en_hlg_pl:*
37:	en_hlg_pr:*
38:	en_hlg_sy:*
39:	en_hlg_un:*
40:	en_hlg_vt:*
41:	en_hgo_om:*
	en_hgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1365	100.0	1365	6	AR078181	Sequence
2	1365	100.0	1365	6	AR208487	Sequence
3	1365	100.0	1365	6	E15382	Bacillus sp
4	1365	100.0	2408	1	AB010272	Bacillus
5	1365	100.0	2408	6	AR078182	Sequence
6	1365	100.0	2408	6	AR208488	Sequence
7	1365	100.0	2408	6	E15383	Sequence
8	755.8	55.4	11318	1	AE007686	Bacillus sp
9	617.2	45.2	1449	1	AX344229	Sequence
10	578.6	42.4	2007	1	BSSACB	X02730 Bacillus su
11	578.6	42.4	5831	12	AF047518	Cloning v
12	578.6	42.4	5842	12	AF004910	Cloning v
13	578.6	42.4	5846	12	XX017500	Cloning vec
14	578.6	42.4	5941	12	SYNPRLC	Cloning vec
15	578.6	42.4	6347	12	SYNPRLA	Cloning vec
16	578.6	42.4	6349	12	AF047519	Cloning v
17	578.6	42.4	6798	12	SYNPRLB	Cloning vec
18	578.6	42.4	9297	12	AF048702	Cloning vec
19	578.6	42.4	88063	1	BS294043	Site-spec
20	578.6	42.4	209510	1	BSU0B018	B. subtilis
21	577	42.3	10317	6	AR162197	Bacillus su
22	565.8	41.5	1422	1	BSU34874	Sequence
23	542.8	39.8	2350	1	BASACB	X34874 Bacillus st
24	542.8	39.8	11612	12	CV080929	u80929 Cloning vec
25	542.8	39.8	16011	12	XXU09128	Cloning vec
26	542.8	39.8	16297	12	CV075991	Cloning vec
27	542.8	39.8	18772	12	AF133437	Cloning v
28	542.8	39.8	19500	12	CV075992	Cloning vec
29	542.8	39.8	22530	12	AB020028	Transform
30	542.8	39.8	107872	2	AC013485	AC013485 Trypanoso
31	542.8	39.8	110000	2	AC058788	Continuation (3 of
32	542.8	39.8	113316	2	AC007866	AC007866 Trypanoso
33	542.8	39.8	122526	2	AC012647	AC012647 Trypanoso
34	542.8	39.8	133928	2	AC091089	AC091089 Oryza sat
35	542.8	39.8	133928	2	AC008146	AC008146 Trypanoso
36	542.8	39.8	146349	2	AC007863	AC007863 Trypanoso
37	542.8	39.8	151443	2	AC007865	AC007865 Trypanoso
38	542.8	39.8	154413	2	AC007862	AC007862 Trypanoso
39	542.8	39.8	155204	2	AC007926	AC007926 Trypanoso
40	542.8	39.8	156128	2	AC113265	AC113265 Mus muscu
41	542.8	39.8	156709	2	AC013484	AC013484 Trypanoso
42	542.8	39.8	156754	2	AC008031	AC008031 Trypanoso
43	542.8	39.8	157774	2	AC013353	AC013353 Trypanoso
44	542.8	39.8	159717	2	AC008368	AC008368 Trypanoso
45	542.8	39.8	162249	2	AC009259	AC009259 Trypanoso

ALIGNMENTS

RESULT 1				
AR078181				
LOCUS	AR078181	1365 bp	DNA	linear PAT 31-AUG-2000
DEFINITION	Sequence 4 from patent US 5962297.			
ACCESSION	AR078181			
VERSION	AR078181.1	GI:10004927		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1365)			
TITLE	Tsusakl,K., Kubota,M. and Chaen,H.			
JOURNAL	Polypeptides having .beta.-actofuranosidase activity			
FEATURES	Patent: US 5962297 A 4 05-OCT-1999;			
	Location/Qualifiers			

```

source          1. 1365
                /organism="unknown"
BASE COUNT      459 a      266 c      283 g      357 t
ORIGIN

Query Match      100.0%; Score 1365; DB 6; Length 1365;
Best Local Similarity 100.0%; Pred. No. 1.7e-305;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAACAGCGGGGACTACAGAGAGACGTATGTTTTGCCCATATTAACGGCCTGACATG 60
1 ATGAACAGCGGGGACTACAGAGAGACGTATGTTTTGCCCATATTAACGGCCTGACATG 60
61 CTAATAATTCAGAGCAACAAAACAGTCCCAATTTAAAGTCCCTCAATTCATGATCA 120
61 CTAATAATTCAGAGCAACAAAACAGTCCCAATTTAAAGTCCCTCAATTCATGATCA 120
121 GCAATCAAAAACATTTGATTCGGCAAAAGGCTATGATAGTCAAGCACTTAATAGATT 180
121 GCAATCAAAAACATTTGATTCGGCAAAAGGCTATGATAGTCAAGCACTTAATAGATT 180
121 GCAATCAAAAACATTTGATTCGGCAAAAGGCTATGATAGTCAAGCACTTAATAGATT 180
181 GATGATAGGATAGTGGCCACTGCAAAAACGCTGATGCTAGCGGCAATTAATGATGA 240
181 GATGATAGGATAGTGGCCACTGCAAAAACGCTGATGCTAGCGGCAATTAATGATGA 240
241 TATCAGATCGTCCGCTTTAGCAGGTGACCCAAAAACAGTATGATCTCCACTTCAT 300
241 TATCAGATCGTCCGCTTTAGCAGGTGACCCAAAAACAGTATGATCTCCACTTCAT 300
301 TTATCTCATCAAAAAGTGGGTATCATGATTTGACAGCTGCAAAAATGCTGGAAGATGA 360
301 TTATCTCATCAAAAAGTGGGTATCATGATTTGACAGCTGCAAAAATGCTGGAAGATGA 360
361 TTTGAAGATATGATTAATTTGTTCCAAATGATCCGTATCTTAATATCAACACAGGAG 420
361 TTTGAAGATATGATTAATTTGTTCCAAATGATCCGTATCTTAATATCAACACAGGAG 420
421 TGGTCAGGTTCTGCTACTTTAAACAAAGATGCGCAAGTCCGTTTATTCATACAGATTAC 480
421 TGGTCAGGTTCTGCTACTTTAAACAAAGATGCGCAAGTCCGTTTATTCATACAGATTAC 480
481 TCAGGTAATTCCTGAAGATGATGGAACCGGTGCTGTAACCAATCATTTCAACTGCTCAA 540
481 TCAGGTAATTCCTGAAGATGATGGAACCGGTGCTGTAACCAATCATTTCAACTGCTCAA 540
541 GTAACCTTATCCACCGCGATGACAGTACACTTAAGTGGATGAGATCTGATCATTA 600
541 GTAACCTTATCCACCGCGATGACAGTACACTTAAGTGGATGAGATCTGATCATTA 600
601 TCTGCTTTGATGCGGAGACGCTACAGTTTATCAAAATATTCACGAATTTATCGATGA 660
601 TCTGCTTTGATGCGGAGACGCTACAGTTTATCAAAATATTCACGAATTTATCGATGA 660
661 GGCAGTGGATTTACAGGTATACCATACTTTAAAGAGCCCTCATATGTAAGTAAG 720
661 GGCAGTGGATTTACAGGTATACCATACTTTAAAGAGCCCTCATATGTAAGTAAG 720
721 GGCCTAATAATATCTGTCTTTGAAGGAAATCTGGAACAACAGATGGTATGAAGCCAT 780
721 GGCCTAATAATATCTGTCTTTGAAGGAAATCTGGAACAACAGATGGTATGAAGCCAT 780
781 CAGTCTTCAATTAATTAAGCTTACTATGCGGAGAGAGCTCTTCTCCAGAAATGAAAA 840
781 CAGTCTTCAATTAATTAAGCTTACTATGCGGAGAGAGCTCTTCTCCAGAAATGAAAA 840
841 AATAAAGCTTCAAAAGTCTTAAACAAATTCCTTTAGCGAATGCTCATTTAGGC 900
841 AATAAAGCTTCAAAAGTCTTAAACAAATTCCTTTAGCGAATGCTCATTTAGGC 900
901 ATTGTTGAATGGCGGATGACTATACAGTGAAGAGTGTATGAAACCATTAATGCTCATCA 960
901 ATTGTTGAATGGCGGATGACTATACAGTGAAGAGTGTATGAAACCATTAATGCTCATCA 960

```

```

1021 AACACAGTAGACAGATGAAGTGCAGACGGCCAAATATATTTAAATGAATTAATGATAT 1020
1021 AACACAGTAGACAGATGAAGTGCAGACGGCCAAATATATTTAAATGAATTAATGATAT 1020
1021 CTAATTCAGGATTTCAAGAGATTCCAAATATGACGATGATGATTAATGACAGCAAGATGTT 1080
1021 CTAATTCAGGATTTCAAGAGATTCCAAATATGACGATGATGATTAATGACAGCAAGATGTT 1080
1081 TATATGCTAGAGCGCGGAGCGGACATCTCTTAATATGAGCCCAACACCGGTAATGAATCT 1140
1081 TATATGCTAGAGCGCGGAGCGGACATCTCTTAATATGAGCCCAACACCGGTAATGAATCT 1140
1141 GCACCTGTAATTAACATGAATCTTGACCCCTGCTGATCTCACACACACTTACTCTATTGC 1200
1141 GCACCTGTAATTAACATGAATCTTGACCCCTGCTGATCTCACACACACTTACTCTATTGC 1200
1201 GGTATTCGGCACCCTGGAAGGTATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1260
1201 GGTATTCGGCACCCTGGAAGGTATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1260
1261 TTCTATCCAGACATCACTCTCACCCTGCGGAGCAAGCTTGAGGTTAATATTAAGGCTCT 1320
1261 TTCTATCCAGACATCACTCTCACCCTGCGGAGCAAGCTTGAGGTTAATATTAAGGCTCT 1320
1321 GACACATCTGAGAGAGAAAATAGTTCCGAGACAAGACAATTTCCCA 1365
1321 GACACATCTGAGAGAGAAAATAGTTCCGAGACAAGACAATTTCCCA 1365

RESULT 2
AR208487      1365 bp      DNA      linear      PAT 20-JUN-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT      459 a      266 c      283 g      357 t
ORIGIN

Query Match      100.0%; Score 1365; DB 6; Length 1365;
Best Local Similarity 100.0%; Pred. No. 1.7e-305;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAACAGCGGGGACTACAGAGAGACGTATGTTTTGCCCATATTAACGGCCTGACATG 60
1 ATGAACAGCGGGGACTACAGAGAGACGTATGTTTTGCCCATATTAACGGCCTGACATG 60
61 CTAATAATTCAGAGCAACAAAACAGTCCCAATTTAAAGTCCCTCAATTCATGATCA 120
61 CTAATAATTCAGAGCAACAAAACAGTCCCAATTTAAAGTCCCTCAATTCATGATCA 120
121 GCAATCAAAAACATTTGATTCGGCAAAAGGCTATGATAGTCAAGCACTTAATAGATT 180
121 GCAATCAAAAACATTTGATTCGGCAAAAGGCTATGATAGTCAAGCACTTAATAGATT 180
121 GCAATCAAAAACATTTGATTCGGCAAAAGGCTATGATAGTCAAGCACTTAATAGATT 180
181 GATGATAGGATAGTGGCCACTGCAAAAACGCTGATGCTAGCGGCAATTAATGATGA 240
181 GATGATAGGATAGTGGCCACTGCAAAAACGCTGATGCTAGCGGCAATTAATGATGA 240
241 TATCAGATCGTCCGCTTTAGCAGGTGACCCAAAAACAGTATGATCTCCACTTCAT 300
241 TATCAGATCGTCCGCTTTAGCAGGTGACCCAAAAACAGTATGATCTCCACTTCAT 300

```

OY	301	TTATCTTACAAAAGTGGGATGATCATCGATTGAGACGCTGGAAAAATCTGGAAACAGTA	360
Db	301	TTATCTTACAAAAGTGGGATGATCATCGATTGAGACGCTGGAAAAATCTGGAAACAGTA	360
OY	361	TTTGAAGATATGAGATMAATTTGTTCCAATGATCCGTATCTTAAATATCAACACAGGAG	420
Db	361	TTTGAAGATATGAGATMAATTTGTTCCAATGATCCGTATCTTAAATATCAACACAGGAG	420
OY	421	TGCTCAGGTTCTGCTACTTTAAACCAAGATGGCCAAAGTCGTTTATCTATACAGATTAC	480
Db	421	TGCTCAGGTTCTGCTACTTTAAACCAAGATGGCCAAAGTCGTTTATCTATACAGATTAC	480
OY	481	TCAGGTAATCCTGAAGATGATGAGAACCGGTCGTGGTAACCAATATTC CAATGATGCAAA	540
Db	481	TCAGGTAATCCTGAAGATGATGAGAACCGGTCGTGGTAACCAATATTC CAATGATGCAAA	540
OY	541	GTAAACTTATCCACCGCGGATGACACTACCTTAAAGTCGATGAGATCTGATCATATAA	600
Db	541	GTAAACTTATCCACCGCGGATGACACTACCTTAAAGTCGATGAGATCTGATCATATAA	600
OY	601	TCTGCTCTTGATGGGAGAGCGGTACAGTTTACAAATATTCACCAATTTATGATGAA	660
Db	601	TCTGCTCTTGATGGGAGAGCGGTACAGTTTACAAATATTCACCAATTTATGATGAA	660
OY	661	GGCAAGTGGATTTCAGGTATTAACCATACCTTTAAGAGACCCCTACTATGTTGAAGATAAG	720
Db	661	GGCAAGTGGATTTCAGGTATTAACCATACCTTTAAGAGACCCCTACTATGTTGAAGATAAG	720
OY	721	GGCCATAAATATCTGTCCTTTGACACCAAAATCTGGAACACAGATGTTATCAAGCGAT	780
Db	721	GGCCATAAATATCTGTCCTTTGACACCAAAATCTGGAACACAGATGTTATCAAGCGAT	780
OY	781	CAGTCCTTTCAAATATAAAGCTTACTATGAGGCGGAATGACGTCCTTTCCAGATGAAAAA	840
Db	781	CAGTCCTTTCAAATATAAAGCTTACTATGAGGCGGAATGACGTCCTTTCCAGATGAAAAA	840
OY	841	AATAAATGCTTCAAAAGTCTTAAAAACAATGCTTTTGAACGAATGTCATTAAGGC	900
Db	841	AATAAATGCTTCAAAAGTCTTAAAAACAATGCTTTTGAACGAATGTCATTAAGGC	900
OY	901	ATTGTTGAATTTGGCGATGACTATACAGTGAAGAAAGGTTATGAACCAATTAAGTCGATCA	960
Db	901	ATTGTTGAATTTGGCGATGACTATACAGTGAAGAAAGGTTATGAACCAATTAAGTCGATCA	960
OY	961	AACACAGTAGAGATGAAGTGGAAAGCGGCCCAATATATTTAAATGAATTAATAATGCTAT	1020
Db	961	AACACAGTAGAGATGAAGTGGAAAGCGGCCCAATATATTTAAATGAATTAATAATGCTAT	1020
OY	1021	CTATTCCAGGATTCAGAGGATCCAAATATGACGATGATGAATTAACGACAAAGATGTT	1080
Db	1021	CTATTCCAGGATTCAGAGGATCCAAATATGACGATGATGAATTAACGACAAAGATGTT	1080
OY	1081	TATATGCTAGAGGCGCGAGGAGTCTTAAATGAGCCCAACAAACCCGATTAATGAAACT	1140
Db	1081	TATATGCTAGAGGCGCGAGGAGTCTTAAATGAGCCCAACAAACCCGATTAATGAAACT	1140
OY	1141	GGACTTGTATTGAACATGAATCTTGACCTGTGATCTCACACACACTTACTCTCTATGCG	1200
Db	1141	GGACTTGTATTGAACATGAATCTTGACCTGTGATCTCACACACACTTACTCTCTATGCG	1200
OY	1201	GGTATCCCGCACCCGGAAGGTTATATGTGTACTCACAAAGTTATATGCGAATGAGAGC	1260
Db	1201	GGTATCCCGCACCCGGAAGGTTATATGTGTACTCACAAAGTTATATGCGAATGAGAGC	1260
OY	1261	TTCTATCCAGAACATCACTCTCACCTGCGGAGCAAGCTTGCGGTTAATATTAAGAGTCT	1320
Db	1261	TTCTATCCAGAACATCACTCTCACCTGCGGAGCAAGCTTGCGGTTAATATTAAGAGTCT	1320
OY	1321	GACACATCTGGAGGAGAAAATAGTCCGACAGCAAGACATTCCTCCA 1365	
Db	1321	GACACATCTGGAGGAGAAAATAGTCCGACAGCAAGACATTCCTCCA 1365	

RESULT 3					
LOCUS	EI5382	1365 bp	DNA	linear	PAT 28-JUL-1999
DEFINITION	Bacillus sp. gene for beta-fructofuranosidase.				
ACCESSION	EI5382				
VERSION	EI5382.1 GI:5710065				
KEYWORDS	JP 199806586-A/1.				
SOURCE	Bacillus sp.				
ORGANISM	Bacillus sp.				
REFERENCE	1. (bases 1 to 1365)				
AUTHORS	Tsusaki K., Kubota M. and Chaen H.				
TITLE	POLYPEPTIDE HAVING BETA-FRUCTOFURANOSIDASE ACTIVITY				
JOURNAL	Patent; JP 199806586-A 1 10-MAR-1998;				
	HAYASHIBARA BIOCHEM LAB INC				
COMMENT	OS Bacillus sp. PN JP 199806586-A/1 PD 10-MAR-1998 PF 09-JUN-1997 JP 1997164875 PR 10-JUN-1996 JP 96P 170630 PI TSUSAKI KEIJI, KUBOTA MICHIO, CHAEN HIROTO PC C12N15/09,C07H21/04,C07K14/32,C12N1/21,C12N9/10,C12N9/24, PC (C12N15/09), PC C12R1(07),(C12N1/21,C12R1:19),(C12N9/24,C12R1:19); CC Strandedness: Double; CC Topology: linear; FH Key Location/Qualifiers FT source 1..1365 /organism='Bacillus sp.' FT FT /strain='V230' FT mat_peptide 1..1365 /product='beta-fructofuranosidase'. FT location/Qualifiers 1..1365 /organism='Bacillus sp.' /db_xref='taxon:1409'				
BASE COUNT	459 a 266 c 283 g 357 t				
ORIGIN					
Query Match:	100.0%; Score 1365; DB 6; Length 1365;				
Best Local Similarity	100.0%; Pred. NO. 1.7e-305;				
Matches 1365; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
OY	1 ATGAACAGCGGGGACTTACAGAAGACGTATTGGCCCATTTATACAGCGCTGCATCG 60				
DB	1 ATGAACAGCGGGGACTTACAGAAGACGTATTGGCCCATTTATACAGCGCTGCATCG 60				
OY	61 CTAAATAATTCAGCAACAACAACAGTCTCTCAATTTAAAGTGCCCTCAATTCATGCATCA 120				
DB	61 CTAAATAATTCAGCAACAACAACAGTCTCTCAATTTAAAGTGCCCTCAATTCATGCATCA 120				
OY	121 GCATTCAAAAACATTGATTTGGGCCAAAABGGATGTAAGTAGTCAGGCACCTTAATAGATTTA 180				
DB	121 GCATTCAAAAACATTGATTTGGGCCAAAABGGATGTAAGTAGTCAGGCACCTTAATAGATTTA 180				
OY	181 GATGTATGAGTAGTACGTGSCCATCTGCAAAACCCTGATGTTACTGCGCGCAAATTCATGATGA 240				
DB	181 GATGTATGAGTAGTACGTGSCCATCTGCAAAACCCTGATGTTACTGCGCGCAAATTCATGATGA 240				
OY	241 TATCACATCGTCTCCGCTTTAGCAGGTGACCCAAAAAACAGTATGATTAATCTCCACTTCAT 300				
DB	241 TATCACATCGTCTCCGCTTTAGCAGGTGACCCAAAAAACAGTATGATTAATCTCCACTTCAT 300				
OY	301 TTATTTCTTCAAAAAGCGGTGATTCATGCAATGTGCACAGCTGGAAAAATGCTGGAAGAGTA 360				
DB	301 TTATTTCTTCAAAAAGCGGTGATTCATGCAATGTGCACAGCTGGAAAAATGCTGGAAGAGTA 360				
OY	361 TTTGAGATGATGAGATAAATTTGTCCAATATATCGTATCTTAAATATCAAAACACAGAG 420				
DB	361 TTTGAGATGATGAGATAAATTTGTCCAATATATCGTATCTTAAATATCAAAACACAGAG 420				
OY	421 TGTCAGGTCTGCTACTTTAAACAAGAATGGCCAAAGTCGCTTTTAATTCATATACAGATTAC 480				

```

Db      421 TGGTCAGGTTCTGCTACTTTAAACAAAGATGGCCAAAGTCGGTTATTCATATACAGATTAC 480
Oy      481 TCAGGTAATCCCTGGAAGATGGTGAACCGGTGCTGGTAACCAATATCTTCAACTGCTCAA 540
Db      481 TCAGGTAATCCCTGGAAGATGGTGAACCGGTGCTGGTAACCAATATCTTCAACTGCTCAA 540
Oy      541 GTAACCTTATCCAGCCGAGTACAGTACATTAAGTGAATGATGATCTGATCTGATCA 600
Db      541 GTAACCTTATCCAGCCGAGTACAGTACATTAAGTGAATGATGATCTGATCTGATCA 600
Oy      601 TCTGCTTTGATGCGGAGAGCGGTACAGTTTATCAAAATATTCAGCAATTTATCGATGA 660
Db      601 TCTGCTTTGATGCGGAGAGCGGTACAGTTTATCAAAATATTCAGCAATTTATCGATGA 660
Oy      661 GGCAGTGGATTTTCCAGGTATACCACTTAAAGAGACCCCTCATATGTAAGATAG 720
Db      661 GGCAGTGGATTTTCCAGGTATACCACTTAAAGAGACCCCTCATATGTAAGATAG 720
Oy      721 GGCATTAATATCTTGTCTTTGAGCGAATACCTGGAACAAGATGTTATCAAGCGAT 780
Db      721 GGCATTAATATCTTGTCTTTGAGCGAATACCTGGAACAAGATGTTATCAAGCGAT 780
Oy      781 CAGCTTTTCAATTAATTAAGCTTACTATAGGCGGAAGTACGCTCTTCCAGATGA 840
Db      781 CAGCTTTTCAATTAATTAAGCTTACTATAGGCGGAAGTACGCTCTTCCAGATGA 840
Oy      841 AATTAACGCTTCAAAAGTCTTAAACAAATATGCTTTTACGCAATGCTGATTAAGC 900
Db      841 AATTAACGCTTCAAAAGTCTTAAACAAATATGCTTTTACGCAATGCTGATTAAGC 900
Oy      901 ATTGTTGAATGGCGGATGACTATACAGTGAAGAGTATATGAACCATTAAGTCGATCA 960
Db      901 ATTGTTGAATGGCGGATGACTATACAGTGAAGAGTATATGAACCATTAAGTCGATCA 960
Oy      961 AACACAGTACAGATGAGTACGAGCGCCCAATATATTAATTAATTAATTAATGAT 1020
Db      961 AACACAGTACAGATGAGTACGAGCGCCCAATATATTAATTAATTAATTAATGAT 1020
Oy      1021 CTATTCAGGATTCAGAGAGATCCAAATAGCAGATGAGTAATTAAGCAAAAGTGT 1080
Db      1021 CTATTCAGGATTCAGAGAGATCCAAATAGCAGATGAGTAATTAAGCAAAAGTGT 1080
Oy      1081 TATATGCTAGGAGCGGAGCGGAGCTCTTAAATGGCCACACACCCGATTAATGAACT 1140
Db      1081 TATATGCTAGGAGCGGAGCGGAGCTCTTAAATGGCCACACACCCGATTAATGAACT 1140
Oy      1141 GGACTTGTATTAACATGATCTTGAACCTGCTGATCTACACACACTTACTCTCAT 1200
Db      1141 GGACTTGTATTAACATGATCTTGAACCTGCTGATCTACACACACTTACTCTCAT 1200
Oy      1201 GGTATCCGCGACCTGGAAGTATATATGTAATGTAATGTAATGTAATGTAATG 1260
Db      1201 GGTATCCGCGACCTGGAAGTATATATGTAATGTAATGTAATGTAATGTAATG 1260
Oy      1261 TTTCTATCAGAAATCACTCTCAGCTGCGGAGCAAGCTTGGGTTAATTAAGGCT 1320
Db      1261 TTTCTATCAGAAATCACTCTCAGCTGCGGAGCAAGCTTGGGTTAATTAAGGCT 1320
Oy      1321 GACACATCTGAGAGGAGAAATAGTTCGCGACAAAGCAATTC 1365
Db      1321 GACACATCTGAGAGGAGAAATAGTTCGCGACAAAGCAATTC 1365

```

RESULT 4

```

AB010272      2408 bp      DNA      linear      BCT 07-AUG-1998
LOCUS      Bacillus sp. gene for Beta-fructofuranosidase, complete cds.
DEFINITION      AB010272
ACCESSION      AB010272.1 GI:3399690
VERSION      Beta-fructofuranosidase
KEYWORDS      Bacillus sp. (strain:V230) DNA.
ORGANISM      Bacillus sp.

```

```

REFERENCE      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
AUTHORS      1 (sites)
TITLE      Kurimoto, M.
JOURNAL      Cloning and sequencing of beta-fructofuranosidase gene from
REFERENCE      Bacillus sp. V230
AUTHORS      Unpublished
TITLE      2 (bases 1 to 2408)
JOURNAL      Tsusaki, K.
JOURNAL      Submitted (08-JAN-1998) Keiji Tsusaki, Hayashibara Biochemical
JOURNAL      Laboratories, Inc., Amase Institute, 7-7 Amase-minami machi,
JOURNAL      Okayama, Okayama 700, Japan (E-mail: amasehepo.harenet.or.jp,
JOURNAL      Tel:086-231-6731, Fax:086-231-6738)
FEATURES      Location/Qualifiers
source      1..2408
            /organism="Bacillus sp."
            /strain="V230"
            /db_xref="taxon:1409"
            /note="Isolated from soil"
            361..1824
            /codon_start=1
            /transl_table=11
            /product="Beta-fructofuranosidase"
            /protein_id="BAA32083.1"
            /db_xref="GI:3399691"
            /translation="MNFRILAKKAAVPTTALIVGADPHIFPAQNNSGYKEDYF
            AHTRADMLKIPGQNSPOKRYPOFNAIXINISAKRYDSSGNLIDIVDPSLON
            ADGTRANHGHTIYSALAGDRNSDDTDLHLRYOKGDTSIDSKNAGRVEDMDKRV
            PNDPLKQTOEWSGATLTLDGQVRLTPTTYSYSGPEGGGAGAQI1STAVQNLSDP
            DAATLVKVDGSDHKSVPDGGDTVYONIOEFLDEKMTLSGNHTLRDPHYVEDKHKY
            LVEFANTGTTDGOESDFNNKAYGSDVFEFONKNTLQSPKQI1SLNGLALGIY
            ELADYTVKSWKPLVASNTVADEVERANTIFRNMNKVYLFIDRSKMTSPGINDKDV
            YMLGPGHSLNGPMPINETGIVLNMDPADIDITHTYSHCIIPIPEGNNVLTISMTN
            RGFYRHHSHLRDKGVNIKSDTSGCGNSSGGQGF"
BASE COUNT      777 a 492 c 496 g 643 t
ORIGIN
Query Match      100.0%; Score 1365; DB 1; Length 2408;
Best Local Similarity 100.0%; Pred. No. 1,66-305;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      1 ATGAACAGCGGGAGCTACAGAGAGAGATGTTTGGCCATATTAACGCGCTGACATG 60
Db      457 ATGAACAGCGGGAGCTACAGAGAGAGATGTTTGGCCATATTAACGCGCTGACATG 516
Oy      61 CTAATAATTCAGGACACAAACAAAGCTCTCAATTAAGTCCCTCATTCATGATCA 120
Db      517 CTAATAATTCAGGACACAAACAAAGCTCTCAATTAAGTCCCTCATTCATGATCA 576
Oy      121 GCAATCAAAACATTTGATTCGCAAAAGGATATGATAGTACAGCACTTAATGATTTA 180
Db      577 GCAATCAAAACATTTGATTCGCAAAAGGATATGATAGTACAGCACTTAATGATTTA 636
Oy      181 GATGTATGGGATAGCTGGCCACTGCAAAACGCTGATGCTGCGGCAAAATATCATGA 240
Db      637 GATGTATGGGATAGCTGGCCACTGCAAAACGCTGATGCTGCGGCAAAATATCATGA 696
Oy      241 TATCAATCGCTCGCTTACAGGATGACCCCAAAACAGATGATACCTCAT 300
Db      697 TATCAATCGCTCGCTTACAGGATGACCCCAAAACAGATGATACCTCAT 756
Oy      301 TTATCTATCAAAAGTGGTGATATCATGATTTGACACCTGGAATAATGCTGGAAGATA 360
Db      757 TTATCTATCAAAAGTGGTGATATCATGATTTGACACCTGGAATAATGCTGGAAGATA 816
Oy      361 TTTGAAGATATGATTAATTTGTTCCAAATGATCGTATCTTAATTAATCAACACAGAG 420
Db      817 TTTGAAGATATGATTAATTTGTTCCAAATGATCGTATCTTAATTAATCAACACAGAG 876
Oy      421 TGGTCAGGTTCTGCTACTTTAAACCAAGATGGCCAGTCCGTTATTTCTATACAGATTAC 480
Db      877 TGGTCAGGTTCTGCTACTTTAAACCAAGATGGCCAGTCCGTTATTTCTATACAGATTAC 936

```


QY 481 TCAGCTAATCCTGAGATGCTGGAGCCGGCTCTGGTAAACAATCATTTCAACTGCTCAA 540
DB TCAGCTAATCCTGAGATGCTGGAGCCGGCTCTGGTAAACAATCATTTCAACTGCTCAA 540
QY 937 TCAGCTAATCCTGAGATGCTGGAGCCGGCTCTGGTAAACAATCATTTCAACTGCTCAA 540
DB 541 GTAACCTATCCCGACCGGATGACACTTAAAGTCGATGGATCTGATCATATA 600
DB 997 GTAACCTATCCCGACCGGATGACACTTAAAGTCGATGGATCTGATCATATA 1056
QY 601 TCTGCTTTGATGGGAGAGCGGTACAGTTATCAAAATATTCACCAATTTATGATGAA 660
DB 1057 TCTGCTTTGATGGGAGAGCGGTACAGTTATCAAAATATTCACCAATTTATGATGAA 1116
QY 661 GCGAAGTGGATTTGAGGATGATACCATCTTTAAGAGACCTCACTATGTTGAAGATAG 720
DB 1117 GCGAAGTGGATTTGAGGATGATACCATCTTTAAGAGACCTCACTATGTTGAAGATAG 1176
QY 721 GCGCATTAATATCTTGTCTTTGAGAGCGAATCTGGAACAACAGATGTTATCAAGCGAT 780
DB 1177 GCGCATTAATATCTTGTCTTTGAGAGCGAATCTGGAACAACAGATGTTATCAAGCGAT 1236
QY 781 CAGTCTTCAATTAATTAAGTCTTACTATGCGGAGAGTACGCTCTTCCAGATGAAAA 840
DB 1237 CAGTCTTCAATTAATTAAGTCTTACTATGCGGAGAGTACGCTCTTCCAGATGAAAA 1296
QY 841 AATAAAGCTCTTCAAGTCTTAAAAAACAAATTCCTTTAGCGAATGCTGCATTAAGC 900
DB 1297 AATAAAGCTCTTCAAGTCTTAAAAAACAAATTCCTTTAGCGAATGCTGCATTAAGC 1356
QY 901 ATTGTTGAATTTGGCGGATGATATGACATGAAATGTTATGAACCATTTAGTGCATCA 960
DB 1357 ATTGTTGAATTTGGCGGATGATATGACATGAAATGTTATGAACCATTTAGTGCATCA 1416
QY 961 AACACAGTACGATGATGAGTGAAGCGCGCAATATATTTAAATGAATTAATGAATGAT 1020
DB 1417 AACACAGTACGATGATGAGTGAAGCGCGCAATATATTTAAATGAATTAATGAATGAT 1476
QY 1021 CTATTCACGATTTCAAGAGATCCAAATTAAGACGATGATGAATTAAGCAACAGATGT 1080
DB 1477 CTATTCACGATTTCAAGAGATCCAAATTAAGACGATGATGAATTAAGCAACAGATGT 1536
QY 1081 TATATGCTAAGGCGCGGAGGAGCTCTTAATATGCGCCACACACACCCGATTAATGAAT 1140
DB 1537 TATATGCTAAGGCGCGGAGGAGCTCTTAATATGCGCCACACACACCCGATTAATGAAT 1596
QY 1141 GGACTTTGATTAATGAATTAATGACCTGCTGATCTCACACACTTACTCTCATTTGC 1200
DB 1597 GGACTTTGATTAATGAATTAATGACCTGCTGATCTCACACACTTACTCTCATTTGC 1656
QY 1201 GGTATCCGCAACCTGAAAGTAAATATGTTGTTACTCACAGTTATATGACGAATAGAGC 1260
DB 1657 GGTATCCGCAACCTGAAAGTAAATATGTTGTTACTCACAGTTATATGACGAATAGAGC 1716
QY 1261 TTTATCCAGAAATCATCTCTACCTGGGAGCAAGCTTTGGGTTAATATTAAGGGTCT 1320
DB 1717 TTTATCCAGAAATCATCTCTACCTGGGAGCAAGCTTTGGGTTAATATTAAGGGTCT 1776
QY 1321 GACACATCTGGAGAGAAATAGTCCGAGACAAGACAATTTCCCA 1365
DB 1777 GACACATCTGGAGAGAAATAGTCCGAGACAAGACAATTTCCCA 1821

RESULT 5
AR078182
LOCUS AR078182 2408 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 5 from patent US 5962297.
ACCESSION AR078182
VERSION AR078182.1 GI:10004928
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2408)
AUTHORS Tsusaki,K., Kubota,M. and Chaeen,H.

TITLE Polypeptides having .beta.-fructofuranosidase activity
JOURNAL Patent: US 5962297-A 5 05-Oct-1999;
FEATURES Location/Qualifiers
source 1. 2408
BASE COUNT 777 a 492 c 496 g 643 t
ORIGIN
Query Match 100.0%; Score 1365; DB 6; Length 2408;
Best Local Similarity 100.0%; Pred. No. 1.6e-305;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACACCGGGGATGATGACAGAAAGATGATGTTTGGCCATATATACAGCGCTGACATG 60
DB 457 ATGAACACCGGGGATGATGACAGAAAGATGATGTTTGGCCATATATACAGCGCTGACATG 516
QY 61 CTAATAATTCAGAGCAACAACAAAGTCCCAATTTAAAGTGCCTCAATTCATATGATCA 120
DB 517 CTAATAATTCAGAGCAACAACAAAGTCCCAATTTAAAGTGCCTCAATTCATATGATCA 576
QY 121 GCAATCAAAACATGATGATGCGCAAAAGGATATGATGATGATGATGATGATGATGATGAT 180
DB 577 GCAATCAAAACATGATGATGCGCAAAAGGATATGATGATGATGATGATGATGATGATGAT 636
QY 181 GATGTATGGGATGATGATGCGCAAAAGGATGATGATGATGATGATGATGATGATGATGAT 240
DB 637 GATGTATGGGATGATGATGCGCAAAAGGATGATGATGATGATGATGATGATGATGATGAT 696
QY 241 TATCATGCTGCTCCGCTTTTACAGGTGACCCCAAAACAGTATGATGATGATGATGATGAT 300
DB 697 TATCATGCTGCTCCGCTTTTACAGGTGACCCCAAAACAGTATGATGATGATGATGATGAT 756
QY 301 TTATCTCTCAAAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 757 TTATCTCTCAAAAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
QY 361 TTTGAAGATATGATGATTAATTTGTTCCAAATGATCCGATCTTAATTAATCAAAACAGAG 420
DB 817 TTTGAAGATATGATGATTAATTTGTTCCAAATGATCCGATCTTAATTAATCAAAACAGAG 876
QY 421 TGGTCAAGTCTGCTACTTTTAACCAAGATGCGCAAGTCCGTTATTTCTATACAGATTAC 480
DB 877 TGGTCAAGTCTGCTACTTTTAACCAAGATGCGCAAGTCCGTTATTTCTATACAGATTAC 936
QY 481 TCAGCTAATCCTGAGATGCTGGAGCCGGTGTGTAACCAATATTTCAAACTGCTCAA 540
DB 937 TCAGCTAATCCTGAGATGCTGGAGCCGGTGTGTAACCAATATTTCAAACTGCTCAA 996
QY 541 GTAACCTATCCCGACCGGATGACACTTAAAGTCGATGGATCTGATCATATA 600
DB 997 GTAACCTATCCCGACCGGATGACACTTAAAGTCGATGGATCTGATCATATA 1056
QY 601 TCTGCTTTGATGGGAGAGCGGTACAGTTATCAAAATATTCACCAATTTATGATGAA 660
DB 1057 TCTGCTTTGATGGGAGAGCGGTACAGTTATCAAAATATTCACCAATTTATGATGAA 1116
QY 661 GCGAAGTGGATTTGAGGATGATACCATCTTTAAGAGACCTCACTATGTTGAAGATAG 720
DB 1117 GCGAAGTGGATTTGAGGATGATACCATCTTTAAGAGACCTCACTATGTTGAAGATAG 1176
QY 721 GCGCATTAATATCTTGTCTTTGAGAGCGAATCTGGAACAACAGATGTTATCAAGCGAT 780
DB 1177 GCGCATTAATATCTTGTCTTTGAGAGCGAATCTGGAACAACAGATGTTATCAAGCGAT 1236
QY 781 CAGTCTTCAATTAATTAAGTCTTACTATGCGGAGAGTACGCTCTTCCAGATGAAAA 840
DB 1237 CAGTCTTCAATTAATTAAGTCTTACTATGCGGAGAGTACGCTCTTCCAGATGAAAA 1296
QY 841 AATAAAGCTCTTCAAGTCTTAAAAAACAAATTCCTTTAGCGAATGCTGCATTAAGC 900
DB 1297 AATAAAGCTCTTCAAGTCTTAAAAAACAAATTCCTTTAGCGAATGCTGCATTAAGC 1356
QY 901 ATTGTTGAATTTGGCGGATGATATGACATGAAATGTTATGAACCATTTAGTGCATCA 960

Db	1337	ATTGTGTAATTTGGCCGATGACTATACAGTGAAGAAAGTGTATGAAACCATTAAGTGCATCA	1416
Qy	961	AACACAGTAGCAGATGAAGTGAACGGCCCAATATATTTAAATGAATTAATTAATGCTAT	1020
Db	1417	AACACAGTAGCAGATGAAGTGAACGGCCCAATATATTTAAATGAATTAATTAATGCTAT	1476
Qy	1021	CTATTTCACGAGTTTAAAGAGATCCAAAATGACGAGTGTGATTAATTAACGACAAAGATGT	1080
Db	1477	CTATTTCACGAGTTTAAAGAGATCCAAAATGACGAGTGTGATTAATTAACGACAAAGATGT	1536
Qy	1081	TATATGCTAGAGGCCCGGAGCGACTCCCTTAAATGAGCCACACAAACCGGATTAATGAAGCT	1140
Db	1537	TATATGCTAGAGGCCCGGAGCGACTCCCTTAAATGAGCCACACAAACCGGATTAATGAAGCT	1596
Qy	1141	GGACTTGTATTGAACATGTAATCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGC	1200
Db	1597	GGACTTGTATTGAACATGTAATCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGC	1656
Qy	1201	GGTATCCCGCACCCTGAAGGTAATTAATGTGTACTCACAGTATTAATGACGAATAGAGGC	1260
Db	1657	GGTATCCCGCACCCTGAAGGTAATTAATGTGTACTCACAGTATTAATGACGAATAGAGGC	1716
Qy	1261	TTCTATCCAGAACATCTCTCACCTGGCGGACAAAGCTTGGGGTTAATATTAAGGATCT	1320
Db	1717	TTCTATCCAGAACATCTCTCACCTGGCGGACAAAGCTTGGGGTTAATATTAAGGATCT	1776
Qy	1321	GACACATCTGAGGAGAGAAAATAGTTCCGGACAGACAAATTTCCCA	1365
Db	1777	GACACATCTGAGGAGAGAAAATAGTTCCGGACAGACAAATTTCCCA	1821
RESULT 6			
AR208488			
LOCUS	AR208488	2408 bp	DNA
DEFINITION	Sequence 5 from patent US 6383769.		linear
ACCESSION	AR208488		
VERSION	AR208488.1	GI:21509655	
KEYWORDS			
ORGANISM	Unknown.		
SOURCE	Unknown.		
REFERENCE	1 (bases 1 to 2408)		
AUTHORS	Tsutsaki, K., Kubota, M. and Chaeu, H.		
TITLE	Polypeptides having .beta.-fluorofuranosidase activity		
JOURNAL	Patent: US 6383769-A 5 07-May-2002;		
FEATURES	Location/Qualifiers		
source	1..2408		
BASE COUNT	777 a 492 c 496 g 643 t		
ORIGIN			
Query Match	100.0%; Score 1365; DB 6; Length 2408;		
Best Local Similarity	100.0%; Pred. No. 1 6e-305;		
Matches 1365; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGAACAGCGGGGACTACAAAGGAAGACTATGGTTTGGCCCATATTACACGGCGTGACATG	60
Db	457	ATGAACAGCGGGGACTACAAAGGAAGACTATGGTTTGGCCCATATTACACGGCGTGACATG	516
Qy	61	CTAAATAATTCAGAGCAACAACAGACGTCCTCAATTTAAAGTGGCTCAATTCATGATCA	120
Db	517	CTAAATAATTCAGAGCAACAACAGACGTCCTCAATTTAAAGTGGCTCAATTCATGATCA	576
Qy	121	GCAATCAAAAAACATTTGATTTGGGCAAAAAGGGATATGATTAAGTCAGGCACTTAATGATTA	180
Db	577	GCAATCAAAAAACATTTGATTTGGGCAAAAAGGGATATGATTAAGTCAGGCACTTAATGATTA	636
Qy	181	GATGTATGGATAGCTGGCCACTCAAAAGCGTATGATGTAATGCGGCAAAATATCATGGA	240
Db	637	GATGTATGGATAGCTGGCCACTCAAAAGCGTATGATGTAATGCGGCAAAATATCATGGA	696
Qy	241	TATCATCATCGTCTCCGCTTTAGCAGGTTGACCCAAAACAGTATGATGATCTCCATCTCAT	300

Db	697	TATCATCGTCTCCGCTTTAGCGAGTGACCCAAAACAGTATGATCTCCACTTCAT	756
QY	301	TTATTCATCAAAAAGTCGTGATACATCATGATTGCACGTGGAAAAATGCTGGAAGATA	360
Db	757	TTATTCATCAAAAAGTCGTGATACATCATGATGCACGTGGAAAAATGCTGGAAGATA	816
QY	361	TTTGACGATATGATAAATTTTGTCCAATGATCCGATATCTTAAATATCAAACACAGAG	420
Db	817	TTTGACGATATGATAAATTTTGTCCAATGATCCGATATCTTAAATATCAAACACAGAG	876
QY	421	TGTCAGTCTCGACTTTTACCACAAAGTGGCCAAAGTCGTTATTTCTATACAGATTAC	480
Db	877	TGTCAGTCTCTGCTACTTTTACCACAAAGTGGCCAAAGTCGTTATTTCTATACAGATTAC	936
QY	481	TCAGTAAATCCGTAAGATGATGGAACCGGTCCTGTTAACCAATCATTTCAACTGCTCAA	540
Db	937	TCAGTAAATCCGTAAGATGATGGAACCGGTCCTGTTAACCAATCATTTCAACTGCTCAA	996
QY	541	GTAACCTTATCCACGCCGATGACGCTTACACTTAAAGTCATGAGATATCTGATCATPAA	600
Db	997	GTAACCTTATCCACGCCGATGACGCTTACACTTAAAGTCATGAGATATCTGATCATPAA	1056
QY	601	TCTGCTTTGATGGGGGACGGATACAGTTATCAAAATATTCAGCAATTTATCGATGAA	660
Db	1057	TCTGCTTTGATGGGGGACGGATACAGTTATCAAAATATTCAGCAATTTATCGATGAA	1116
QY	661	GGCAAGTGCATTTGACGATACCATATCTTAAAGACCCCTCACTATGTTGAAGATPAG	720
Db	1117	GGCAAGTGCATTTGACGATACCATATCTTAAAGACCCCTCACTATGTTGAAGATPAG	1176
QY	721	GGCCATAATATCTTGTCTTTGAAGCGATATCTGSAACACAGATGCTTATCAAGCGAT	780
Db	1177	GGCCATAATATCTTGTCTTTGAAGCGATATCTGSAACACAGATGCTTATCAAGCGAT	1236
QY	781	CAGCTTTTCATATTAAGCTTACTATGSGGAAGTGACGCTCTTCCAGAAATGAAGAAA	840
Db	1237	CAGCTTTTCATATTAAGCTTACTATGSGGAAGTGACGCTCTTCCAGAAATGAAGAAA	1296
QY	841	AATAAAGCTGCTTCAAGAGCTCTTAAAAAACAAATTCCTTTTACGGAATGGTGATTAAGG	900
Db	1297	AATAAAGCTGCTTCAAGAGCTCTTAAAAAACAAATTCCTTTTACGGAATGGTGATTAAGG	1356
QY	901	ATTGTTGATTTGGCCGATGACTATPACGTGAAAAGTGTATGAACCATTAAGTCGCATCA	960
Db	1357	ATTGTTGATTTGGCCGATGACTATPACGTGAAAAGTGTATGAACCATTAAGTCGCATCA	1416
QY	961	AACACAGTAGAGATGAAGTGAACGGCCCATATTTTAAATGAATAAATAAATGAT	1020
Db	1417	AACACAGTAGAGATGAAGTGAACGGCCCATATTTTAAATGAATAAATAAATGAT	1476
QY	1021	CTATTCACGAGATTAAGAGGATCCAAAATGACAGAGTGTGAATTTAACAGCAAAAGATGT	1080
Db	1477	CTATTCACGAGATTAAGAGGATCCAAAATGACAGAGTGTGAATTTAACAGCAAAAGATGT	1536
QY	1081	TATATGCTAGGCCCCGAGGCGACTCCTTAAATGGCCACACAACCCGATTAATGAACCT	1140
Db	1537	TATATGCTAGGCCCCGAGGCGACTCCTTAAATGGCCACACAACCCGATTAATGAACCT	1596
QY	1141	GGACTTGTATGAACATGATATCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGC	1200
Db	1597	GGACTTGTATGAACATGATATCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGC	1656
QY	1201	GGTATCCCGCACCCCTGAAGATTAATATGTTACTACACAACTTATATGACCAATPAGAGGC	1260
Db	1657	GGTATCCCGCACCCCTGAAGATTAATATGTTACTACACAACTTATATGACCAATPAGAGGC	1716
QY	1261	TTTCTATCCAGAACATCACTCTCACCTGGCGGACAGCTTGGGCTTAAATATTTAAAGGCTCT	1320
Db	1717	TTTCTATCCAGAACATCACTCTCACCTGGCGGACAGCTTGGGCTTAAATATTTAAAGGCTCT	1776
QY	1321	GACACATCTGGAGGAGAAATAGTTCGGGACAGGACAAATTTCCCA	1365

Db 1777 GACACATCTGGAGGAGAAATAGTTCCGACAGACACATTCACA 1821

RESULT 7
LOCUS E15383 2408 bp DNA linear PAT 28-JUL-1999
DEFINITION *Bacillus* sp. gene for beta-fructofuranosidase, complete cds.
ACCESSION E15383.1 GI:5710066
VERSION JP 199806586-A/2.
KEYWORDS
SOURCE *Bacillus* sp.
ORGANISM *Bacillus* sp.
Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
REFERENCE 1 (bases 1 to 2408)
AUTHORS Tsusaki, K., Kubota, M. and Chaen, H.
TITLE POLYPEPTIDE HAVING BETA-FRUCTOFURANOSIDASE ACTIVITY
JOURNAL Patent: JP 199806586-A 2 10-MAR-1998;
HAYASHIBARA BIOCHEM LAB INC

COMMENT
OS *Bacillus* sp.
PN JP 199806586-A/2
PD 10-MAR-1998
PE 09-JUN-1997 JP 1997164875
PR 10-JUN-1996 JP 96P 170630
PI TSUSAKI KEIJI, KUBOTA MICHIO, CHAEN HIROTO
PC C12N15/09, C07H21/04, C07K14/32, C12N1/21, C12N9/10, C12N9/24, PC
(C12N15/09,
PC C12R1:07), (C12N1/21, C12R1:19), (C12N9/24, C12R1:19); CC
strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1. 2408
FT /organism='Bacillus sp.'
FT /strain='V230'
FT CDS 361..1824
FT /product='beta-fructofuranosidase' FT
sig_peptide 361..456.
Location/Qualifiers
1. 2408
/organism='Bacillus sp.'
/db_xref='taxon:1409'
BASE COUNT 777 a 492 c 496 g 643 t
ORIGIN

Query Match 100.0%; Score 1365; DB 6; Length 2408;
Best Local Similarity 100.0%; Pred. No 1.6e-305;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACAGCGGGGACTACAGAGAGACTATGTTTGGCCATATTACAGCGCTGACATG 60
Db 457 ATGAACAGCGGGGACTACAGAGAGACTATGTTTGGCCATATTACAGCGCTGACATG 516
QY 61 CTAAATTTCCAGACACAAACAGTCTTAATTTAAAGTCCCTCAATTAAATGATCA 120
Db 517 CTAAATTTCCAGACACAAACAGTCTTAATTTAAAGTCCCTCAATTAAATGATCA 576
QY 121 GCAATCAAAAACATTTGATTTGGCAAAAAGGATGATAAGTCAGGCACTTAATGATTTA 180
Db 577 GCAATCAAAAACATTTGATTTGGCAAAAAGGATGATAAGTCAGGCACTTAATGATTTA 636
QY 181 GATGTATGGATAGTACCTGCGCACTGCAAAAAGCTGATGTTAGCGCAAAATATCATGCA 240
Db 637 GATGTATGGATAGTACCTGCGCACTGCAAAAAGCTGATGTTAGCGCAAAATATCATGCA 696
QY 241 TATACATTCGTCCTGCTTTTATGACAGGTGACCCCAAAAACATGATGATCTCAGCTTCAT 300
Db 697 TATACATTCGTCCTGCTTTTATGACAGGTGACCCCAAAAACATGATGATCTCAGCTTCAT 756
QY 301 TTATTTATCAAAAAGTGGTATGATCATGATTTGACAGCTGGAAAAATGCTGAAGAGTA 360
Db 757 TTATTTATCAAAAAGTGGTATGATCATGATTTGACAGCTGGAAAAATGCTGAAGAGTA 816
QY 361 TTGAGATATGATGATTAATTTGTTCCAAATGATCCGATCTTAAATATCAACACAGAG 420

Db 817 TTTGAGATATGATGATTAATTTGTTCCAAATGATCCGATCTTAAATATCAACACAGAG 876
QY 421 TGGTCAGGTCTGCTACTTTTAAACCAAGATGGCCAGCGCTTATCTATACAGATTAC 480
Db 877 TGGTCAGGTCTGCTACTTTTAAACCAAGATGGCCAGCGCTTATCTATACAGATTAC 936
QY 481 TCAGGTAAATCCGTAAGATGTGGAACCGGTCTGTAACCAATATCAATTCAGTCTCAA 540
Db 937 TCAGGTAAATCCGTAAGATGTGGAACCGGTCTGTAACCAATATCAATTCAGTCTCAA 996
QY 541 GTAACCTTATCCAGCGCGGATGACACTACACTTAAAGTCGATGATCTGATCATPAA 600
Db 997 GTAACCTTATCCAGCGCGGATGACACTACACTTAAAGTCGATGATCTGATCATPAA 1056
QY 601 TCTGCTTTGATGGCGGAGAGAGTACAGTATATCAAAATATTCAGCAATTTATGATGAA 660
Db 1057 TCTGCTTTGATGGCGGAGAGAGTACAGTATATCAAAATATTCAGCAATTTATGATGAA 1116
QY 661 GGCAAGTGGATTTGAGTGTATACCATCTTAAAGAGACCCCTCATGTTGAAGATPAG 720
Db 1117 GGCAAGTGGATTTGAGTGTATACCATCTTAAAGAGACCCCTCATGTTGAAGATPAG 1176
QY 721 GGCCATTAATATCTTGTCTTTGAAGCGAATACGGAACACAGATGTTATCAAGCGCAT 780
Db 1177 GGCCATTAATATCTTGTCTTTGAAGCGAATACGGAACACAGATGTTATCAAGCGCAT 1236
QY 781 CAGTCTTCATTAATTAAGCTTACTATGCGGAAGTGACGCTCTTCCAGAAATGAAAA 840
Db 1237 CAGTCTTCATTAATTAAGCTTACTATGCGGAAGTGACGCTCTTCCAGAAATGAAAA 1296
QY 841 AATAAAGCTCTTCAAAAGCTCTTAAAAACAATTCCTTTAGGCAATGCTGAAGGC 900
Db 1297 AATAAAGCTCTTCAAAAGCTCTTAAAAACAATTCCTTTAGGCAATGCTGAAGGC 1356
QY 901 ATTGTTGAATTTGGCCGATGACTATACAGTGAAGAAGTGTATGAACAATTAAGTCGATCA 960
Db 1357 ATTGTTGAATTTGGCCGATGACTATACAGTGAAGAAGTGTATGAACAATTAAGTCGATCA 1416
QY 961 AANCACAGTACAGATGAGAGTGGAGCGCCCAATATTTAAATGAATTAATGATGAT 1020
Db 1417 AANCACAGTACAGATGAGAGTGGAGCGCCCAATATTTAAATGAATTAATGATGAT 1476
QY 1021 CTATTCACGAGATTCAGAGAGATCCAAATATGACGAGTGTGAATTAACGACAAAAGATGT 1080
Db 1477 CTATTCACGAGATTCAGAGAGATCCAAATATGACGAGTGTGAATTAACGACAAAAGATGT 1536
QY 1081 TATATGCTAAGGCGCCGAGGAGACTCTTAAATGCGCCACACAAACCCGATTAATGAAT 1140
Db 1537 TATATGCTAAGGCGCCGAGGAGACTCTTAAATGCGCCACACAAACCCGATTAATGAAT 1596
QY 1141 GGACTGTGATTTGAACATTAATCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGC 1200
Db 1597 GGACTGTGATTTGAACATTAATCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGC 1656
QY 1201 GGTATCCCGCACCTGAGAGTGAATATGTTGATGACACAAATTTATGACGAATAGAGGC 1260
Db 1657 GGTATCCCGCACCTGAGAGTGAATATGTTGATGACACAAATTTATGACGAATAGAGGC 1716
QY 1261 TTTATCCAGAACATCACTCTACCTGCGGGACAAAGCTTGGGGTTAATATTAAAGGCTCT 1320
Db 1717 TTTATCCAGAACATCACTCTACCTGCGGGACAAAGCTTGGGGTTAATATTAAAGGCTCT 1776
QY 1321 GACACATCTGGAGGAGAAATAGTTCCGGAACAGGACATTTCCCA 1365
Db 1777 GACACATCTGGAGGAGAAATAGTTCCGGAACAGGACATTTCCCA 1821

RESULT 8
LOCUS AE007686 11318 bp DNA linear BCT 27-JUL-2001
DEFINITION *Clostridium acetobutylicum* ATCC824 section 174 of 356 of the
complete genome.

ACCESSION AE007686 AE001437
VERSION AE007686.1 GI:15024739
KEYWORDS
SOURCE
ORGANISM Clostridium acetobutylicum.
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium
REFERENCE 1 (bases 1 to 11318)
Nolling, J., Breton, G., Omelchenko, M. V., Markarova, K. S., Zeng, O.,
Gibson, R., Lee, H. M., Dubois, J., Oiu, D., Hitti, J., Wolf, Y. I.,
Tatusov, R. L., Sabathe, F., Doucette-Stamm, L., Soucaille, P.,
Daly, M. J., Bennett, G. N., Koonin, E. V. and Smith, D. R.
Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum
JOURNAL J. Bacteriol. 183 (16), 4823-4838 (2001)
MEDLINE 21359325
PUBMED 11466286
TITLE 2 (bases 1 to 11318)
REFERENCE Childress, D., Zeng, O. and Smith, D. R.
AUTHORS Direct Submission
TITLE Submitted (24-JUN-2001) GTC Sequencing Center Production,
JOURNAL Flushing and Bioinformatics Teams, Genome Therapeutics Corp., 100
Beaver Street, Waltham, MA 02453-8443, USA
FEATURES
source
1. 11318
/organism="Clostridium acetobutylicum"
/strain="ATCC 824"
/db_xref="ATCC:824"
/db_xref="taxon:1488"
89. 790
/gene="CAC1770"
89. 790
/gene="CAC1770"
/codon_start=1
/transl_table=11
/product="RNA polymerase sigma factor, SigK-like"
/protein_id="AAK79735.1"
/db_xref="GI:15024740"
/translation="MLENRSLEDRLQKAKESKELNRLIOEKPEIASLLOKTKGYL
OYGDDELSTGIMAFEEAVESEYDRNGKELSPARKVITLRAIDYRRKQKRNIVYLA
GVNSDENIDIRPIENKAMEYKKNKSINERIRLELEPKRLREMGIEFVULVNASPK
OEKLKRYKNIANIVIHMODILNKLIRKRLPIKEITEKNITINRKLKERGIYITALV
IAEIGDRLIREYLE"
801. 1544
/gene="CAC1771"
801. 1544
/gene="CAC1771"
/codon_start=1
/transl_table=11
/product="uncharacterized protein, ykri B subtilis
homolog"
/protein_id="AAK79736.1"
/db_xref="GI:15024741"
/translation="MKGIVIELEKYKLIILEKDFEIDKLNTRAKIGOEINIKSKK
ISTIKALREFLAATALLIEPIINSLAYGMFTREYVVDINPNNNKRVAMIEIYNF
GNIIRLOAANKGVSIIVKMRNLKFKATNVINNEIRTAQDKFIDSGKONTIYTTA
SSSISIDISIDSELEHYKENKINARPMIVLNDIDYKSKQYGIPIDKTILLNKVI
KNPSYKPYDINKKSIDELINIANQENY"
1682. 3151
/gene="CAC1772"
/note="sacB"
1682. 3151
/gene="CAC1772"
/codon_start=1
/transl_table=11
/product="Levanucrase"
/protein_id="AAK79737.1"
/db_xref="GI:15024742"
/translation="MNKLKIVKICLIGSMISGIIITOOTFASTNDMNYKENTYVSHIT
RYNSKIPMEQNDLKFRYPOENASTIKNIASAKYDKNGNLDIDVDSMPLQWGDGT
VANHGHIYFALAGDKPRNODDTIYMPYQKIGENSIDSKNACKVFRDSKYVANDP
YKLTOTEMSSATUTSDGOVRLFTYDFSEVAKDGDTDASQVITTTQVNLSDPSNT
INIDVSDHKSVEFDGNGTITVQVQFIDECKWSSGDNHTLRDPHYVEDNGKXYLVE

ANTGTNDYOGDTSLNKAFYGRSOSFEKTEKQDLIDTNKKHDSLANGALGIIELN
NDYTLKKEKMPPLIASNTVTEIERANGFKMGKMYLFTDSGSKMTNGISSKIIYAL
GFSSNSLGPYKPKLNKGLVNLNDLPDTLFTYSHFAVPQTNKNVITSYITNRGM
YSDHSSFPAPSPFLINIKGTSTVSNISLDOGLTIDNY"
3249. 4793
/gene="CAC1773"
/note="sacB"
3249. 4793
/gene="CAC1773"
/codon_start=1
/transl_table=11
/product="Levanase"
/protein_id="AAK79738.1"
/db_xref="GI:15024743"
/translation="MKRRKLKILFIETFIILIVLSLTIQNNSTIKSTTKKONY
REYVHTVPNKKRNDPQRPVPEDEYHYVLYLNQDNGNCTEMROATSDIATWRE
GCIPKTYTNGDITWSSGFVYDAONTGFEFGKALVALVTPPSASMDQEOFLSTGR
GRTFRPYSNOPIIIPNCTKDFRDKIIMDERKNKNWVLAEGTIGFYESYLNKNQW
TGDFETNNIGIVECPDIEMQSDNGNKKWILGTSANCKVSGEPYVAVWNGDGKKE
IADISTPKMLDYGEDWYAAVTFESENOSKILKRYALAMNNMDYANNPTIQRNFG
MDSIVREITLSKQKNDNTYSLSKRIKIENTITSIDQKQISVGLKHLVQGAQYL
PDTINWDARKWQIRLRESLDKRRHIDVCFTEBKYSYVNRATYGNPDKSKKYESA
PFDINNKVHLRLIFVDRVSEVFIIDGRITYSNEVFRPEKGTILFSIKKAVFKNI
TIKHID"
4860. 6146
/gene="CAC1774"
/note="sacB2"
4860. 6146
/gene="CAC1774"
/codon_start=1
/transl_table=11
/product="Levanucrase"
/protein_id="AAK79739.1"
/db_xref="GI:15024744"
/translation="MKTRKTYKMISLMLIATITPELLIRHNTGYTSIHSROAON
FKCTKENTAPINPNEFKLTAENLWMDPTWLVKDGSLAVANGIKVIFALPASNVAM
NKRHDVAGISYECSTDENWYKGLAVNEDALGRQWASAIIDENGMOQEFYTAG
RKGEAVTFEQLRATKTFESINVDKGVAHITCSKHQVILPEDGVYIOTMOQAKPIT
SERPVFEEDPKTKDYLIPEGNGKGVKIEKKRPENIGDKLERKHIAPRGVENGVY
GIAVAOKKDLRFRLPLLEAVGVNOOLERPOIWKVKNKYLIPTISHKFTYAGQNLG
VDGIXGCNGLSNBNRYPLNNGVLTINPTNDPQYTSWLYVSHDVLSFINETHFNG
QLRYGTFFAPLQIISLGVKSKITIGLGEGVYPAH"
6560. 8305
/gene="CAC1775"
6560. 8305
/gene="CAC1775"
/codon_start=1
/transl_table=11
/product="Predicted membrane protein"
/protein_id="AAK79740.1"
/db_xref="GI:15024745"
/translation="MKIKGIIAISMSLITGSISYVYKASSNNKRLONTYKEROKIK
LNKSNKSLVNTDNEKLSLDDLNKKNKVIPEFNKTIINNKIKLPIITSGDNSKTK
GNSVASAGDSNNAVPAKLDPTYLANEBSLKNBITTNTQIAWIMSDPADGSLSA
VGGPSPAYIILGLDNGKPYQTFNADPDFLYOVYTPSKGKSNIVGSIHVIPPEY
OAFQGTPOSTIDVTYVNAVDESKMOTAICTGTSDYMNKLISDQSGVAVWVYV
DKKVVYINKPSNAGCINNTITITPVLSSNTSVSRMLMGNKNDEAMLSQONTI
ILDKTYENNKQFOETYPKNQOMEPETADGSVITLNNYPELRFQVPTBLOVYL
PDTNDPNNASTHRSQYCGLNKTKAEKRLNTAGODVLYIYSPSLIOTDPIILDEH
LTVGDPAVLIYDSTTAVADNYITGPTTYSSTANINVTNVPATAVKDIYASLNDVAR
VSDENYKVMAPGESFWRDSGOQCINIGIDVYVNGNTNQLMCTWRFVQAGVDPIS
IIPGLHFDYVEAGD"
8693. 9346
/gene="CAC1776"
8693. 9346
/gene="CAC1776"
/codon_start=1
/transl_table=11
/product="HAD superfamily hydrolase"
/protein_id="AAK79741.1"
/db_xref="GI:15024746"
/translation="MDRIIFEDCYGLIDTRKPYVEDAISNIKNNILNSKDAVINIYVY
EDRLMGETYIPIYDKLIKQVLEYCDMELVTNVFSKSYDEIISVHKRLKPKREVESLE

KHKYELALMSVHDIIDYNIDALGNFNNVLSEDIHAYPPOLKFEKVEERLE
LKORHCHIAWMDIIVPCSLGMRKIWRKNKKGKHEPEVEVHTDEVINLI*
9598. 10521
/gene="CAC1777"
9598. 10521
/gene="CAC1777"
/codon_start=1
/transl_table=1
/product="Nudix (Mult-like) hydrolase"
/protein_id="AAK79742.1"
/db_xref="GI:15024747"
/translation="MRDNKKNRNGLTKEKFLDRYPGDERPSNTVMDLFTVDVY
IEGDPKALKILIKRGDHYMGCAVPGGFVNIINGLSACRYRELKEETNVNVE
EOLTFEDVDYRDMRIVISVYMALADKISITPKAGDDADADAFVTKKEFISKGF
GIDKIDYINILITDDETRIGLVTEKFEKNGVITIKVPSYRLLEMSKEELADHIE
EIDCALBRINKIEYPIAPSRILPKYFTLIRAKVFEAILNLKPLIRANRRRIKKM
VVEREKRTISGRATCTFEDMOTFLDE"
10543. 11193
/gene="CAC1778"
10543. 11193
/gene="CAC1778"
/codon_start=1
/transl_table=1
/product="Amidase from nicotianamide family"
/protein_id="AAK9743.1"
/db_xref="GI:15024748"
/translation="MGKFEKILNMSEIKKFEVDNLDDKDAIVVDVNGFVHE
GALSSPVEGIVDEIVINEKTELGKRIKFLDEHTNOSTERYSAKHOLESDIELI
PELKNELLDSTNIVMIRPKNSYNGFHAGFKWLEPNESQIENYITCCGVEDICSNFA
NTLKTYNOKMMDRIIIPSNVETFDGHDLMKILISLEMOSNGIETVDRVL"
BASE COUNT 4466 a 1373 c 1945 g 3534 t
ORIGIN
Query Match 55.4% Score 755.8; DB 1: Length 11318;
Best Local Similarity 73.4% Pred. No. 1.4e-164;
Matches 980; Conservative 0; Mismatches 352; Indels 3; Gaps 1;
QY 12 GGACTACAGGAGACATGATGTTTGGCCATATTACAGCGCTGACATGCTAAATAATCC 71
DB 1777 GAATTACAAAGAACATATGATGATTTGCGATATTACGCTTATATATGCTAAATAATCC 1836
QY 72 AGGACACAAACACGCTCTCAATTTAAAGTCCCTCAATTCATGATGATGACCAATCAAAA 131
DB 1837 AATGACACAAACATCTTAAATTTAAAGTCCCTCAATTTCAATCATCAATCAATTAATAA 1896
QY 132 CATGATTCGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 191
DB 1897 TATAGCTTCAGCAAAAGGTTATGATAGCAATGTTATGATGATGATGATGATGATGATG 1956
QY 192 TAGCTGGCCACTGCAAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATG 251
DB 1957 TAGCTGGCCACTGCAAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2016
QY 252 CTCCGCTTTAGCAGTACCCAAAAACAGATGATGATGATGATGATGATGATGATGATGATGATG 311
DB 2017 TTTTGGCTTTGGCAGTATCTTAAATCAAGATGATGATGATGATGATGATGATGATGATGATG 2076
QY 312 AAAAGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 371
DB 2077 AAAAGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2136
QY 372 GGATTAATTTGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 431
DB 2137 TGATTAATTTGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2196
QY 432 TGTACTTTTAAACCAAGATGCGCAAGTCGTTTATTCATACAGATTAAGTATGATGATGATG 491
DB 2197 TGTACTTTTAAACCAAGATGCGCAAGTCGTTTATTCATACAGATTAAGTATGATGATGATG 2256
QY 492 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 551
DB 2257 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2316

QY 552 CCAGCCGGATGACAGTACACTTAAAGTGATGATGATGATGATGATGATGATGATGATGATGATG 611
DB 2317 TCAACAGATTCATCAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2376
QY 612 TGGCGAGACGCTACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 671
DB 2377 CGGAGGAAATGGGAAATCTATCAGAACGCTTACAGAAATTAATTAATTAATTAATTAATTAAT 2436
QY 672 TTCAGGCTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 731
DB 2437 TTTTGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2496
QY 732 TCTTGTCTTTGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 791
DB 2497 TCTTGTCTTTGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2556
QY 792 TAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 848
DB 2557 CAATTAAGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2616
QY 849 GCTTCAAGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 908
DB 2617 AATTGATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2676
QY 909 ATTGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 968
DB 2677 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2736
QY 969 ACCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1028
DB 2737 AACGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2796
QY 1029 GATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1088
DB 2797 AGATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2856
QY 1089 AGGCGCGAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1148
DB 2857 AGGTTTGTGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2916
QY 1149 ATTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
DB 2917 ACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2976
QY 1209 GACCCCTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1268
DB 2977 TCAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3036
QY 1269 AGAATCATCTCTACCTGCGGAGCAAGCTTGGGTTAATTAATTAATTAATTAATTAATTAAT 1328
DB 3037 AGATCATCATCTCTAGTTTGGACCAAGTTTCTACTTAATTAATTAATTAATTAATTAATTAAT 3096
QY 1329 TGGAGAGCAAAATAG 1343
DB 3097 TGTATATCAAAATAG 3111
RESULT 9
AX434229
LOCUS AX434229 1449 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 2644 from Patent WO0229113.
ACCESSION AX434229
VERSION AX434229.1 GI:21659037
KEYWORDS
SOURCE
ORGANISM Bacillus licheniformis.
REFERENCE 1. Berka, R. and Clausen, I. G.
METHODS for monitoring multiple gene expression
Patent: WO 0229113-A 2644 11-APR-2002;
Novozymes Biotech, Inc. (US); Novozymes A/S (DK)
FEATURES
Location/Qualifiers

	repeat_unit	337. .370	/translation="MCKGQSYFEGVTPYIF"
	RBS	446. 455	/note="imp. Inverted repeat A "
	CDS	464. 1885	/note="pot. RNA binding site (sacB)"
		464. .1885	/codon_start=1
			/transl_table=1
			/product="levansucrase"
			/protein_id="CA26513.1"
			/db_xref="GI:732568"
			/db_xref="SWISS-PROT: P05655"
			/translation="MNKKRKAQATVLTFTTALLAGCATQAFAKETNOKPYKETYGIS HIRHMDLIQPEQOKNEKYQPEDETSIKINISSAKLDVDSWPLONADGTANYAG YHITFALADQPKNADDTISYMFQKVGETSIDSWRNAGRVKEDSKDANDSILKDOT QEMSGATFTSDGKILRFLPYDSEGKHGKDTLTAAOVNASIDSILNINCEYDKSTF DGGKATYONQOFTIDECNYSGGDHTLRDHYHEDKSHKTLVPEANTGTEDEGGEES LNFNAYIGKSTSFPROESOKLLOSDKRTRELANGALGMIELNDYTLKRYMKPLIAS NKTVIDELERANVEFMNGKMYLFTDSRSKMTIDGITSNDIYMLGYSNSLTGPYKPLN RTGVLKMDLMDPNDVPTFYSHFAVPAQKGNVVITSYTNRGFYADQSTFAFPELLN IKGKTSVVKDSILDEGQLTVNK"
	sig_peptide	464. .550	
	mat_peptide	551. .1882	
	misc_feature	1992. .1928	/product="mature levansucrase (aa 1-444)"
		1892. .1928	/note="region of dyad symmetry"
	repeat_unit	1892. .1904	
			/note="inverted repeat B"
	repeat_unit	1916. .1928	
			/note="inverted repeat B'"
BASE COUNT		720 a 422 c 370 g 495 t	
ORIGIN			

Query Match	42.4%;	Score 578.6;	DB 1;	Length 2007;
Best Local Similarity	67.2%;	Pred. No. 1.6e-123;		
Matches 903; Conservative	0;	Mismatches 389;	Indels 51;	Gaps 4

OY	ATGAAACGCGGGGAGCTCAACAGGAAGCATATGCGTTTGGCCATATTACAGCGCGTCAATG	60
Db	ACGAAACCAAAAGCCATATTAAAGAAACATTAGCGCATTTGCCATATTACACGCCATATG	616
OY	CTAAAAATTCGAGGACACAAACACAGTCCCTCAATTAAAGTCCCTCAATTCAATGATCA	120
Db	CTGCAATCCCTGSAACAGCAAAAATGAAATATCAAGTCCCTGAATTGATTCGTC	676
OY	GCATTCAAAAACATTGGATTGGGCAAAAGGATATGATATAGTCAGCAACTTAATGATTTA	180
Db	ACAATTTAAAAATATCTCTTCGCAAAAAG-----CCTG	709
OY	GATGTATGGGATAGCTGGCCACCTGCAAAAGCGTGATGTACTGGCGCAATTATCATGGA	240
Db	GACGTTGGGACAGCTGGCCATTACAAAAGCGCTACGCGCACTGTCCGAACATACACGGC	769
OY	TATCAATCGTCTCCGCTTTAGCAGGTGACCCAAAAACAAGTATGATCTCCACTTCAT	300
Db	TACCAACATCGCTTTGACATTAGCGGAGATGCTTAAATGCGGATGACACATCATATTAC	829
OY	TTATTCATCAAAAAGTCGGTGAATACATGCTTACAGCTGAGAAATGCTGGAAGAAGTA	360
Db	ATGTTCTATCAAAAAGTCGGGAAACTTCTATTGACAGCTGGAAAAACGCTGGCGGCTC	889
OY	TTTGAAGATATGATAAATTTGTTCCAAATGATCCGATCTTAAATATCAAAACACAGAG	420
Db	TTTAAAGACAGCGACAATTCGATGCAATGATTTCTAATAAAGACCAACACAGAA	949
OY	TGGTCAGGTTCTGTAATTTAACCAAGATGCGCAAGTCGTTATTCGTATACGATTAAC	480
Db	TGGTCAGGTTGAGCCATTTTACATCTGACGGAATAATCCGTTTATTCCTACACGATTC	1009

Oy	481	CCAGGTAATCCTCGAAGATGGTGGAAACCGGTGCTGTGAACCAAAATCATTTCAACGCTCCAA	540
Db	1010	TCGCGTAAACCTTA-----CGGCAACCAAAACCTGACAACTGCACAA	1051
Oy	541	GTAACACTTAATCCAGCCGATGCACCTACCTTAACCTCATGAGATGCATCATAA	600
Db	1052	GTTAAAGCATACGATCAGACAGCTCT---TGAACATCAAGGGTGTAGAGATTTATAA	1108
Oy	601	TCGTCTTTTGAATGGCGGAGCGTACAGTTTATCAAAATATTCGCAATTTATGCATGAA	660
Db	1109	TCATCTTTTGA---CGGTGACGCAAAAAACGATATAAATGACAGCGATTCATGATGAA	1165
Oy	661	GGCAAGTGGATTCAGGTGATACCACTTAATTAAGACCTCCTACTATGTTGAAGATTAAG	720
Db	1166	GGCAACTACACCTAGCGGACAAACCTATAGCTGAGATCTCTACTGTAAGATTA	1225
Oy	721	GGCCATAAATATCTTGCTTTTGAAGCAATACTGGAACAAACAGATGGATTCAAGGGAT	780
Db	1226	GGCCAAATAATCTTAGTTTGAAGCAACACCTGGAACTGAAAGATGGCTACCAAGCGGA	1285
Oy	781	CAGTCTTTCAATTAATTAAGCTTACTATGCGCGAAGTGAAGCTTCTTCCAGATGAAAA	840
Db	1286	GAACTCTTATTAACAAAGCATACTATGAGCAAAAGACATCATCTTCCGCAAGAAAGT	1345
Oy	841	AATAACTGCTTCAAAGTCTTAAAAAACAATTCCTTCTTAGGGAATGTGCAATTAGGC	900
Db	1346	CAAAACACTTCGCAAAAGGATTAACAAAGCGACGGCTAGTTAGCAAAAGCGCGCTCGGT	1405
Oy	901	ATTGTGTGAATGGCGCGATGACTATACAGTGAAGAGTGTATGAACCAATTTAGTGCATCA	960
Db	1406	ATGATTTGAGCTTAACGATGATTACACACTGAAAAAGATGAAGAAACCGCTGATTCATCT	1465
Oy	961	AACACAGTAGCAGATGAAGTGAAGCGGCCAATATATTTAAATGAATTAATTAATGSGTAT	1020
Db	1466	AACACAGTAACAGATGAATGAAAGCGGCGCAAGCCTTAATGAAGCAAGCAATGGTAC	1525
Oy	1021	CTATTCAAGGATTCAGAGATCCAAATAATGACAGTATGGAATTTAACGCAAAAGATGTT	1080
Db	1526	CTGTTACTGACTCCCGCGGCTCAAAAATGACGATGACGGCATTAGCTCTAACGATATT	1585
Oy	1081	TATATGCTAGAGGCGCGGAGCGGACCTCCTTAATGAGCCACACAAACCGATTAATGAACCT	1140
Db	1586	TACATGCTTGGTTTATGTTTCTTAATTTCTTTAACTGGCCATACAAAGCCGCTAACAAAACCT	1645
Oy	1141	GGACTTGATTTGAACATGAATCTTGACCCCTGCTATGTCACACACACTTACTCTCATATGC	1200
Db	1646	GGCCTTGTTGTTAAAAATGATCTTGATCCTAACGATGAACCTTTACTACTACACATTC	1705
Oy	1201	GGTATCCCGCACCTGGAAGTATATATGTGTACTACACAGTATTATAGCAATAGAGGC	1260
Db	1706	GCTGTACTCTCAAGGAAAGAAACAACTCTGTGATTACAACTTATAGCAAAACGAGGA	1765
Oy	1261	TTTATATCAGAAACATCACTCTCACCTGGGGGACAAAGCTTGGGGTAAATATTAAGGGCT	1320
Db	1766	TTTCAAGCAGCAAAATCAACGTTTGGCCGAAGCTTCTGTGAAACATCAAAAGGCAAG	1825
Oy	1321	GACACATCTGGAGGAGAAATAG	1343
Db	1826	AAAACATCTGTTGTCAAAAGACAG	1848
RESULT 11			
AF047518/c		5831 bp	DNA
LOCUS	AF047518		circular SYN 22-JUL-1998
DEFINITION	Cloning vector pEX18cm,		complete sequence.
ACCESSION	AF047518		
VERSION	AF047518.1		GI:2961140
KEYWORDS			
SOURCE	Cloning vector pEX18cm.		
ORGANISM	Cloning vector pEX18cm		
REFERENCE	1 (bases 1 to 5831)		
AUTHORS	Hoang,T.T., Karkhoff-Schweizer,R.R., Kutchina,A.J. and		


```

OY 661 GGCAGTGGATTTTCAGTGTATACCTACTTTTAAGAGACCCCTACTATGTGAAGATAG 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 928 GGCACATACAGTACGGGACACACATACCGTCAGACATCCCTACTAGTAGAATATA 869
OY 721 GGCCTAATATCTTGTCTTTGAAGCGAATCTGCAACACAGATGTTATCAGCGCAT 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 868 GGCACAATATCTGTATTTGACCAACACTGCACTGAAGATGGCTACCAAGGCCAA 809
OY 781 CAGTCTTCATTAATTAAGCTTACTATGGCGGAATGACGCTTCTCCAGATGAAAA 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 808 GATCTTTATTTATTAACAAAGCATACTATGTGCAAAAGACATCTTCTCCGTCAGAAAGT 749
OY 841 AATAAAGCTCTCAAGTCTTAAAAAACAATTTGCTTTAGCGAATGTGCATTAGGC 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 748 CAAAAAGCTTGCAGAAAGGATATAAAAGCAACGGCTGAGTTAGCAAAAGGGCGCTCGGT 689
OY 901 ATTGTGAATTTGGCGGATGACTATACATGAAAGTGTATGAACCATTTAGTCGATCA 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 688 ATGATTTGAGCTAAAGATGATTAACACTGAAAAAAGATGAAAAAGCGCTGATTCATCT 629
OY 961 AACACAGTAGACGATGAGTGAAGCGCCCAATATATTAATTAATTAATTAATGATAT 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 628 AACACAGTAGACGATGAGTGAAGCGCCCAATATATTAATTAATTAATTAATGATAT 569
OY 1021 CTATTACAGGATTCAGAGATCCAAATGACGAGTGAATTAACGACAAAGATGTT 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 568 CTGTCTACTGATCCCGCGGACCAAAATGACGATTTGACGGCATACGCTTAACGATAT 509
OY 1081 TATATGTAAGGCGCGGAGCGGACCTCTTAATGCGCCACACACACCCGATTAATTAAC 1140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 508 TACATCTGTGTTATGTTCTTAATCTTAACGACCATCAACCCGCTGAACAAACT 449
OY 1141 GGACTTGTATTAATGATGATTTGACCCCTGCTGATCTCACACACATTAATCTATGC 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 448 GGCCTGTGTTAAATGATGATTTGATCTTAACGATGATTAACCTTTACTTACACACTTC 389
OY 1201 GGTATCCCGACCTGGAAGTAAATGTTGTTACTTCAAGATTAATGACGATAGAGGC 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 388 GCTGTACTCTCAAGCGCAAGCAACATGTCTGATTAACAGCTATATGACAAACAGAGA 329
OY 1261 TTTATTCAGAAACATCACTCTACCTGCGGGGACGACGCTTGGGTTAATTAAGCGTCT 1320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 329 TTTTACGAGACAAACATCAACGTTTGGCGCTACTTCTGCTGACATCAAAAGGCAAG 269
OY 1321 GACACATCTGGAGGAGAAATAG 1343
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 268 AAAACATCTGTTGCAAGACAG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
XXU17500/ 5846 bp DNA circular SVN 18-DEC-1994
LOCUS Cloning vector pEX100T, complete sequence.
DEFINITION u17500
ACCESSION u17500.1 GI:603175
VERSION 1
KEYWORDS beta-lactamase (bla); lac alpha peptide (laczalpha); levansucrase
(mesab); origin of transfer (ori); transcription terminators;
SOURCE Cloning vector pEX100T.
ORGANISM Artificial sequences; vectors.
REFERENCE 1 (bases 1 to 5846)
AUTHORS Schweizer,H.P. and Hoang,T.T.
TITLE An improved system for gene replacement and xyle fusion analysis in
Pseudomonas aeruginosa
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5846)
AUTHORS Schweizer,H.P.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-1994) Herbert P. Schweizer, Microbiology &
Infectious Diseases, University of Calgary, 3330 Hospital Drive
N.W., Calgary, Alberta T2N 4N1, Canada
FEATURES Location/Qualifiers

```

```

source
1..5846
/organism="Cloning vector pEX100T"
/species="host="Escherichia coli"
/db_xref="taxon:37970"
/plasmid="pEX100T"
/complement(209..1630)
/gene="sacB"
/complement(209..1630)
CDS
/codon_start=1
/transl_table=11
/product="levansucrase"
/protein_id="AA57325.1"
/db_xref="GI:603178"
/transl_table="MSI0HFRVALIPFAFCLEPVAHPETLVKVKAEQDQGVAGV
YHIVFALAGDPRKADDTSTYHETKQVETSIDSKKNNGRPRYKDSKDDANDYIKDOT
QEMSGSATFTSDGKIRLEVTDESGRKHGKOTLTAQVNSASDSSLTINDVEDYKSLF
DGDGRTYQNVQOFLDEBNYSAGDHTLRDPHYVDKHKYLVFPAANTGDEGYGYES
LEFNKAYYKSTSPFROESOKLLOSDKKRTAEALGALGMIELANDYTLKKVMKPLIAS
NTVIDEIRAVNFKNMGKMYLFTOSRGSKMTIDGITSNDIYMLGYVNSLTPGPKPLN
KTGLVTKMDLPDNPFTYSHFAYPQAKGNVITISYINRGFYADKOSTFAPSFLN
IKKRTSYVRSILLEGOLTVNK"
2104..2928
/note="region containing ori"
misc_feature
/complement(2984..3309)
/note="E. coli 5S RNA and terminators"
terminator
/complement(2984..3011)
/note="T2"
terminator
/complement(3018..3143)
/note="T1"
rRNA
/complement(3022..3309)
/gene="5S ribosomal RNA"
gene
/complement(3345..3629)
/gene="laczalpha"
CDS
/complement(3345..3629)
/gene="laczalpha"
/codon_start=1
/transl_table=11
/product="lac alpha peptide"
/protein_id="AA57324.1"
/db_xref="GI:603177"
/transl_table="MTMTTPSSRDNRVYIPGRDNRVMSLAVALVLRDWDENPGVTOIN
RLAHPFASWRNSEARTDRPSOQLRSLNENRAMEFWREDEDFDPDT"
3655..3605
/note="polylinker"
misc_feature
/complement(4786..5646)
/gene="bla"
CDS
/complement(4786..5646)
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA57323.1"
/db_xref="GI:603176"
/transl_table="MSI0HFRVALIPFAFCLEPVAHPETLVKVKAEQDQGVAGV
YHIVFALAGDPRKADDTSTYHETKQVETSIDSKKNNGRPRYKDSKDDANDYIKDOT
QEMSGSATFTSDGKIRLEVTDESGRKHGKOTLTAQVNSASDSSLTINDVEDYKSLF
DGDGRTYQNVQOFLDEBNYSAGDHTLRDPHYVDKHKYLVFPAANTGDEGYGYES
LEFNKAYYKSTSPFROESOKLLOSDKKRTAEALGALGMIELANDYTLKKVMKPLIAS
NTVIDEIRAVNFKNMGKMYLFTOSRGSKMTIDGITSNDIYMLGYVNSLTPGPKPLN
KTGLVTKMDLPDNPFTYSHFAYPQAKGNVITISYINRGFYADKOSTFAPSFLN
IKKRTSYVRSILLEGOLTVNK"
BASE COUNT 1400 a 1395 c 1430 g 1621 t
ORIGIN
Query Match 42.4% Score 578.6 DB 12 Length 5846;
Best Local Similarity 67.2% Pred. No. 1.5e-123;
Matches 903; Conservative 0; Mismatches 389; Indels 51; Gaps 4;
OY 1 ATGAACAGCGGGAGCTACAGGAGACATGTTTGGCCATTTATACAGCGGCTGACATG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1537 ACGAACCAAAAGCCATATAAGAAACATACGCGCATTTCCATATTACGCGCATGATATG 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 61 CTAAAAATTCAGGACACAAACAGTCTCAATTTAAAGTGCTCAATTCATGATCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      1477 CTGCAAAATCCCTGGAACGCAAAAAAATGAAAAATATCAAGTTCCGTGATTCGATTCCTCC
Oy      121  GCATTAACAAAACATGATTCGCAAAAGGGATATGATAGTCAGGCAACTTAATGATTTA 180
Db      1417 ACAATTAATAAATATCTCTTCTGCAAAAG-----CCTG 1385
Oy      181  GATGTATGGATAGTACGCGCACTGCAAAACGCTGATGCTGCGCAATTAATCATGGA 240
Db      1384 GACGTTTGGGACACTGGCCATTACAAAACGCTGACGCACTGTCGCAAACTATCCAGGC 1325
Oy      241  TATCAATGCTCTCCGCTTTAGCAGGTGACCAAAAAACAGTATGATATCCACTTCAT 300
Db      1324 TACCAATCCGCTTTGCAATTTAGCCGAGATCTTAAAAATGGGATGACACATCATTTAC 1265
Oy      301  TTATTTATCAAAAAGTGGGATATACATGATGATGACGCTGGAATAATGCTGGAAGACTA 360
Db      1264 ATGTTCTATCAAAAAGTGGGAAACTTCTATGACGCTGGAATAATGCTGCGCGCTC 1205
Oy      361  TTGAGATATGATTAATTTGTTCCAAATGATCCGATCTTAAATATCAACACAGAG 420
Db      1204 TTTAAGACAGCAGCAAAATGATGCAAAATGATCTATCTTAAAGACCAACACAGAA 1145
Oy      421  TGGTCAGTTCTGCTACTTTAACCAAGATGCCAGTCCGTTATTTCTATACAGATTAC 480
Db      1144 TGGTCAGTTCTGACACATTTACATCTGACGAAAAATCCGTTATTTCTACATGATTTC 1085
Oy      481  TCAGTATCTCTGAAGATGTTGGAACCGGTCGTGTAACCAATCTTCAACGCTCAA 540
Db      1084 TCCGTTAAACATTA-----CGGCAACCAACACTGACACATGCAACAA 1043
Oy      541  GTAACCTATCCAGCGGATGACATACCTTAAGTCAGTGAAGTATGATCATATAA 600
Db      1042 GTTAACGATACAGATGACAGAGCTCT--TTGAACATCAACGGTATGAGATATATA 986
Oy      601  TCTGCTTTGATGGCGGACGATGACATTTATAAATTTACGCAATTTATGATGAA 660
Db      985 TCAATCTTTGA---CGGTGACGGAATAACGTAATAATGATACAGCACTTATCATGATA 929
Oy      661  GGCAAGTGAATTTAGGATATACCATCTTAAGAGACCTCTATGTAAGATATAG 720
Db      928 GGCACATACAGCTAGGCGACACATACGCTGAGAGATCTCTACATGATGAGATATA 869
Oy      721  GGCCATAAATATCTTCTTTGAAGCGAATACTGGAACACAGATGTTATCAAGCGCAT 780
Db      868 GGCACAAATATCTAGTATTTGAAGCAAAACACTGGAACCTGGAACGCTACCAAGCGAA 809
Oy      781  CAGTCTTTCAATATAAAGCTTACTATGGCGGAAGTACGCTTCTTCCACATGAAAAA 840
Db      808 GAATCTTTATTTAAACAAAGCACTATGCGCAAAAGCACATCTTCTCCGTCAGAAAAGT 749
Oy      841  AATAAAGCTTTCAAAAGCTTAAACAAACAAATGCTTCTTGAAGATGTCATTTAGC 900
Db      748 CAATAAATCTTGCAGAAAGCGATTAATAAACGACGCGCTAGATTAGCAAGCGCGCTCTCGT 689
Oy      901  ATTGTGAATTTGGCGGATGACTATACAGTGAAGATGTTATGAACCACTTATGTCGCATCA 960
Db      688 ATGATTTAGATAAAGATGATTAACACTGAAAAAGATGAAAAACCGCTGATTGCACT 629
Oy      961  AACACAGTACAGATGAGTGAAGCGGCAATATATTTAAATGAATTAATGAT 1020
Db      628 AACACAGTACAGATGAGTGAAGCGGCAAGCTCTTTAAATGAAGCGCAAAATGAT 569
Oy      1021 CTATTCAGGATTTAAGAGATCCAAATGAGATGATGATTAACAGCAAGATTT 1080
Db      568 CTGTTCACTGACTCCCGCGGATCAAAAATGACGATGACGCAATTAAGTCTTAACGATTT 509
Oy      1081 TATATGCTAGGCGCGGAGCGACTCTTAATATGCGCCACAAACCGGATTAATGAAT 1140
Db      508 TACATGCTGCTGATGTTTCTTAATTTCTTAACGCGCATACAGCGCGTCAACAAACT 449
Oy      1141 GCAATTTATTAATGATGATCTTGCCTCTGATCTCACACACACTTACTCTCATTTCC 1200

```

```

Db      448  GGCCCTTGTTAAATAATGATCTTGATCTTAACGATGATTAACCTTACTTACTCACACTTC 389
Oy      1201  GGTATCCCGGACCCGAGGATTAATATGTTGATGACACAACTTATGACGAATAGAGGC 1260
Db      388  GCTGATCTTCAAGCAAGAAAGAAACATGCTGATTTACAACTTATGACAAACAGAGA 329
Oy      1261  TTTATCCAGAAACATCTTACCTGCGGACAAAGCTTGGGCTTAATATTAAGAGGCTCT 1320
Db      328  TTTATCCAGAGCAAAACATCAACGTTTGGCTGCTTCCGTCGTAACATCAAAAGCGCAG 269
Oy      1321  GACACATCTGAGAGGAATAATAG 1343
Db      268  AAACATCTGTTGCAAGACAG 246

RESULT 14
SYNPRIC/LOCUS 5941 bp DNA circular SYN 27-APR-1993
DEFINITION Cloning vector (pRL278) for sacB-mediated positive selection for double recombinants in gram-negative bacteria.
ACCESSION L05083
VERSION L05083.1 GI:209132
KEYWORDS Cloning vector.
SOURCE Cloning vector DNA.
ORGANISM unidentified cloning vector
REFERENCE 1 (bases 1 to 5941)
AUTHORS Cai, Y.
TITLE Molecular genetic approaches towards the understanding of heterocyst differentiation and pattern formation in the cyanobacterium Anabaena sp
unpublished (1992)

JOURNAL Black, T.A., Cai, Y. and Wolk, C.P.
REFERENCE 2 (sites)
AUTHORS Spatial expression and autoregulation of hetR, a gene involved in the control of heterocyst development in Anabaena
unpublished (1992)

JOURNAL Location/Qualifiers
FEATURES
source 1..5941
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
complement(473..1894)
/feature="sacB from B.subtilis; ORF"
/codon_start=1
/transl_table=11
/protein_id="AAJ72307.1"
/db_xref="GI:209133"
/transl_table=11
/transl_table="MNIRKFAKQATVITFTALLAGATQAFAKETNOKPYKETYGIS
HITRHMLOIPEOQKNEKYQVPEPDSSTIKNISAKGLEWDSMPLOADGTVAHYG
YHIVFALAGDKNDKSDTSIYMFYKVGSTSIDSKNAGRVKSDKFDANDSILKDOT
OEMSGSATFTSDGKIRLEFYDPSGKHGKOTLTTAAYVVASDSSINNGVEDYKSI
DGDCKTQONNOOPIDEGNSSGDNHITLDPHYVEDKGIKIVFPFANGTENGTCGES
LPKAYVGTSTFPEKESOKLLOSDKRTAELANGALGMLINDYTLKVKMPLAS
NTVTEDEIRANVFNKMGKMYLFTDSRSKMTIDITSNDIYMLGVSLNGPYKPLN
IKGKTSYVKDLSILEOGOLTVNK"
2749..3543
/feature="kanamycin and neomycin resistance determinant
(inp11)"
/codon_start=1
/transl_table=11
/protein_id="AAJ72308.1"
/db_xref="GI:209134"
/transl_table="MIEDGIGHASPAWVERLFGYMAOQTIGCSDAAPRLSAOGR
PVLVTKIDLSALNELODEARLSMLATTPCAAVLVDYEAERDMILLGEVGDIL
LSHLDAPEKVISINADADARRLITDPATCPDHOAKHIERARMEGLVDODDDE
EHQGLAPAEFLARLKARMPDEDELVVTHGDACTENINAVENRFGFTIDCGRLVADRY
ODIALATRIDIAELFEGEWADEFLVLYGIAAPDSORIAFRLIDEP"
3558..4815
/feature="multiple cloning sites"
/feature="misc-feature
misc-feature 4816..5941
/feature="oriV and oriT genes from plasmid pMB1"
BASE COUNT 1419 a 1329 c 1471 g 1722 t

```

ORIGIN

Query Match 42.4%; Score 578.6; DB 12; Length 5941;
 Best Local Similarity 67.2%; Pred. No. 1.5e-123;
 Matches 903; Conservative 0; Mismatches 389; Indels 51; Gaps 4;

```

OY 1 ATGAACAGCGGGGACATACAGGAAGACTGTTTGGCCATATTATACAGCGCGTACATG 60
Db 1801 ACGAACAACAAAGCATATAGGAAGAAATACGCGCATTTCCATATTACAGCGCATGATATG 1742
OY 61 CTAAAAATTCAGAGACACAAACAAAGCTCCATATTAAAGTCCCATTTCAATGATCA 120
Db 1741 CTGCAAAATCCCTGAACAGCAAAAAAATGAAAAATATCAAGTTCCGAAATTCGATCGTCC 1682
OY 121 GCAATCAAAAACATTTGATTCGGCAAAAAGGTATGTAAGTACAGCACTTAATAGATTTA 180
Db 1681 ACANTTAATAATATCTCTTCGCAAAAG-----CCTG 1649
OY 181 GATGATGGATAGCTGGCCACTGCAAAAAGCTGATGTTACTCGGCAATTTATCANTGA 240
Db 1648 GAGGTTTGGACAGCTGGCCACTTACAAAACGCTGCGCACTGTCGCAACTATCAGCGC 1589
OY 241 TATCAGATCGTCCGCTTTAGCAGGTGACCCAAAACAGTATGATCCACTTCAT 300
Db 1588 TACCAATCGTCTTTGATTAAGCCGAGATCTTAAATATGCGATGACACATCGATTTC 1529
OY 301 TTAATCTATCAAAAAGTGGTGTATACATCGATTGACAGCTGGAATAATGCTGAAGATTA 360
Db 1528 ATGTTCTATCAAAAAGTGGGGAACCTCTATTTGACACTGGAATAAGCTGGCGGCTC 1469
OY 361 TTTGAGATATGATTAATTTGTTCCAAATGATCCGTATCTTAATATCAACACAGAG 420
Db 1468 TTTAAAGACAGCGCAAAATTCGATGCAATATCTTATCTAAAGACCAACACMAA 1409
OY 421 TGTGAGGTCTGCTACTTTAAACCAAGTGGCCAAAGCTTTTATATACAGATTAC 480
Db 1408 TGGTCAGGTTTACCCACATTTACATCTGACGAAAAATCCGTTTATCTACAGATTTC 1349
OY 481 TCAGATTAATCTGTAAGATGTGGAACCGTCTGTGTAACCAATCATTTCAACTGCTCAA 540
Db 1348 TCCGCTAAACATTA-----CGCCAAACAAACACATGACAACTGCACAA 1307
OY 541 GTAATCTATCCCAAGCCGATGACACTTAAAGTCGATGGATGATCATATAA 600
Db 1306 GTTAACGATACGATCGATGACAGCTCT--TTGAACATCAACGGGTGAGAGATTATAA 1250
OY 601 TCTGCTTTGATGGCGAGACGATACACTTATCAAAATATTCACAAATTTATCATGATAA 660
Db 1249 TCATCTTTGA---CGGTGACGGAACAAACGATATCAAAATGTACAGCACTTCATCATGAA 1193
OY 661 GGCAGATGATTTGAGTGTATACCATATTAAAGACCTCCTACTATGTTGAAGATAG 720
Db 1192 GGCAGATGATTTGAGTGTATACCATATTAAAGACCTCCTACTATGTTGAAGATAG 1133
OY 721 GGCATTAATATCTGCTTTGGAAGCGAATCTGGAACACAGATGTTATCAGGCGAT 780
Db 1132 GGCACCAAAATCTTAATTTGGAACAAACACTGGAAGTGAAGTGGCTACCAAGCGCAA 1073
OY 781 CAGTCTTCATTAATTAAGCTTACTATGCGGAGTACGCTCTTCTCCAGATGAATAA 840
Db 1072 GAATCTTTATTAACAAAGCATACTATATGCAAAAGCATATCTTTCCCTCCAGAAAGT 1013
OY 841 AATAAAGCTCTTCAAGTCTTAATAAACAAATGCTTCTTAAAGCAATGCTGATTAAGC 900
Db 1012 CAATAAAGCTCTTCAAGGATTAATAAACAAAGCAAGGCTGATTAAGCAAGGCTGCTGCT 953
OY 901 ATTGATGATTTGGCGGATGATATACAGTGAAGATGTTAAGAACCACTTATGTCATCA 960
Db 952 ATGATTTGAGTAAAGCATGATTAACACTGAATAAAGTGAATAAACCCCTGATTCATCT 893
OY 961 AACAGATGATGAAGTGAAGCGGCGCAATATATTAATAATGAATTAATATGATAT 1020
Db 892 AACACAGTAAAGATGAATGAAGCGGCAACGCTTTTAAATGAACGCAAAATGCTAC 833

```

```

OY 1021 CTATTCAGCATTCAGAGCATTCCAAAATGACGAGTATGATTAACGACAAAGATGTT 1080
Db 832 CTGTTCACTGACTCCCGGGATCAAAAATGACGATTAAGCGCATTAAGTCTAAGATAT 773
OY 1081 TATATGCTAGGGCCCGGAGCGACTCTTAAATGSCCACAACCCGATTAATGAAT 1140
Db 772 TACATGCTGTTGTTATTTCTTAATTTCTTAATCTGAGCCCATTAAGCCGCTGAACAAAT 713
OY 1141 GGACTGTGATGAACATGAATCTGACCCCTGATATGACATGACACACCTTACTACTTC 1200
Db 712 GGCTTGTTGTTAAATGATCTTGAATCTTAACGATGTAACCTTTACTACTACACTTC 653
OY 1201 GGTATCCCGCACCTTGAAGTAAATGTTGTTACTACACAGTTATATGACGAATAGAGC 1260
Db 652 GCTGATCTCAAGGAAGAAACATATGCTGATTTAAACCTATATGACAAACAGAGA 593
OY 1261 TTTATCCAGAACATCACTTCACCTGCGGACAAAGCTTGCGGTTAATATTAAGGCTCT 1320
Db 592 TTTACGCGACAAACATCAACGTTTGGCCCAAGCTTCCTGCAACATCAAAAGCGAAG 533
OY 1321 GACACATCTGAGGAGAAATAG 1343
Db 532 AAAACATCTGTTTCAAGACAG 510

```

```

RESULT 15
SYNPLA/C 6347 bp DNA circular SYN 27-APR-1993
DEFINITION Cloning vector (pRL271) for sacB-mediated positive selection for double recombinants in gram-negative bacteria.
ACCESSION L05081
VERSION L05081.1 GI:209125
KEYWORDS cloning vector.
SOURCE Cloning vector DNA.
ORGANISM unidentified cloning vector.
REFERENCE 1 (bases 1 to 6347)
AUTHORS Cal, Y.
TITLE Molecular genetic approaches towards the understanding of heterocyst differentiation and pattern formation in the cyanobacterium Anabaena sp
JOURNAL unpublished (1992)
REFERENCE 2 (sites)
AUTHORS Black, T.A., Cal, Y. and Wolk, C.P.
TITLE Spatial expression and autoregulation of hetr, a gene involved in the control of heterocyst development in Anabaena
JOURNAL unpublished (1992)
FEATURES
source
location/Qualifiers
1..6347
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
complement(473..1894)
/feature="sacB gene from B. subtilis; ORF"
CDS
1..6347
/db_xref="GI:209126"
/translation="MNIRKFAKONATVLTFTTALLAGTQAFKATNOKPKYKTYGIS
HITRHMLOIPEQOKNEKYVPERPDSITIKNISAKLEIWDSPRLNOACTVANYG
YHIVFALAGDKRKNADDTSTYWEIQRKVEISIDSKNAGRFKDSKDFDANDSILKDOT
QEWGSAITFTSDGKIRLEVTDFSGKHGKQTLTAOVNVSAOSSSLININGVEDKSLF
DGDGKYOVNOVOFIDECNYSAGDHTLRDPHYVDKSHKLYLVEPANTGTEGXYGEEB
LNFKAAYKSTSFRESOKLJOSDKRTAEANGALGMLENDYTLKVKMPLIIS
NTVNDIEKRAVFEKMNKRWLYFDSRSGSKMTIDGITSNDIYMLGTVNSLITGPYKPLN
KTGLVLKNDLDPNDVTFTSYFAVPAQKANNVITTSYTNRGTADNOSTFAPSFLN
IKGKTSVVKDSIIIEGQATVNR"
complement(2393..3127)
/feature="erythromycin resistance determinant"
/codon_start=1
/transl_table=11
/protein_id="AA072303.1"
/db_xref="GI:209127"

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:50:32 ; Search time 1977.86 Seconds
(without alignments)
11177.167 Million cell updates/sec

Title: US-09-986-682B-4
Perfect score: 1365
Sequence: 1 ATGACACAGCGGAGCTACAA.....CCGACACAGACAATTCCTCA 1365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359.6	26.3	755	17	B49063
2	309.6	22.7	623	17	A0284271
3	307.4	22.5	654	17	AG158197
4	296.8	21.7	687	17	AG145469
5	292.6	21.4	689	17	AG174787
6	290.8	21.3	653	17	A0311782

7	286.6	21.0	688	17	AG165661	AG165661 Pan trogl
8	284.4	20.8	656	17	AG158505	AG158505 Pan trogl
9	283.4	20.8	678	17	AG019951	AG019951 Homo sapi
10	271	19.9	711	17	AG001039	AG001039 Homo sapi
11	267	19.6	669	17	AG156831	AG156831 Pan trogl
12	264.8	19.4	581	17	AG091204	AG091204 HS_4355_B
13	258.4	18.9	682	17	AG020006	AG020006 Homo sapi
14	250	18.3	934	9	AL575209	AL575209 AL575209
15	248.6	18.2	696	17	AG185716	AG185716 Pan trogl
16	233	17.1	683	17	AG167579	AG167579 Pan trogl
17	228.6	16.7	622	17	AG160823	AG160823 Pan trogl
18	226	16.6	508	17	AG0317303	AG0317303 RPT11-10
19	220.8	16.2	1101	17	CNS00E5C	AL068761 Drosophila
20	215.8	15.8	488	17	AG0315317	AG0315317 RPT11-10
21	213	15.6	687	17	AG145995	AG145995 Pan trogl
22	194.6	14.3	661	17	AG019210	AG019210 Homo sapi
23	194.6	14.3	678	17	AG020005	AG020005 Homo sapi
24	192	14.1	698	17	AG165439	AG165439 Pan trogl
25	187.6	13.7	694	17	AG145903	AG145903 Pan trogl
26	186.6	13.7	385	17	B47667	B47667 RPT11-1A4
27	184.8	13.5	694	17	AG169167	AG169167 Pan trogl
28	183	13.4	682	17	AG165125	AG165125 Pan trogl
29	182.6	13.4	714	17	AG014172	AG014172 Homo sapi
30	179	13.1	674	17	AG165337	AG165337 Pan trogl
31	179	13.1	696	17	AG177482	AG177482 Pan trogl
32	173.6	12.7	729	17	AG014170	AG014170 Homo sapi
33	172.2	12.6	698	17	AG180444	AG180444 Pan trogl
34	170.8	12.5	601	17	P859L	AL390538 Leishmani
35	167.6	12.3	676	17	AG145781	AG145781 Pan trogl
36	167.2	12.2	705	17	AG147097	AG147097 Pan trogl
37	166.8	12.2	722	17	AG014192	AG014192 Homo sapi
38	165	12.1	683	17	AG168295	AG168295 Pan trogl
39	163.2	12.0	935	17	CNS008JC	AL052150 Drosophila
40	162.8	11.9	714	17	AG014191	AG014191 Homo sapi
41	162.4	11.9	672	17	AG149729	AG149729 Pan trogl
42	158.4	11.6	613	17	AG064045	AG064045 927P1-13C
43	154.6	11.3	752	17	AG171011	AG171011 Pan trogl
44	153.6	11.3	704	17	AG168501	AG168501 Pan trogl
45	151	11.1	731	17	AG162359	AG162359 Pan trogl

ALIGNMENTS

RESULT 1
LOCUS B49063 755 bp DNA linear GSS 08-APR-1999
DEFINITION RPT11-4121.TV RPT1-11 Homo sapiens genomic clone RPT1-11-4121, DNA sequence.
ACCESSION B49063
VERSION B49063.1 GI:2601300
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 755)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building
Unpublished (1997)
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are derived from the human BAC library RPT1-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT		
Db	301	GGCGCGGAATGTTTCAAAATGAAGCGAAATGTGACTTGTTCACTGATGTACCGGGTTC	1044	CAAAATGACGAGTGTGAATTAACAGCAAAAGATGTTATATGCTAGGCGCCGAGGCGA	1103	361	AAAATGACGATCGATGCTATTAACACTCAAAAGATTTATACATGCTTGTTATGATCAAA	420	1104	CTCTTAATGAGCCACACAAACCGCATTAATGAAGATGAGCTGTGATTTATGACATGAATCT	1163	421	CTCTTAACCGGCCCTTCAACCGCGCTAAACAAACAGGCGTGTGTGTCGCAAAATGGGCT	480
Db	481	TGATCCAAAGCATGTGACATTCACCTACTCTCACCTGCTGCCAGTCCGCAAGCAAAAGCA	540	1224	TAATGTGCTACTCACACACTTATATGACGAATAGAGCTTCT	1264	541	CAATGTGCTTATCACAAAGCTACATGACAAACAGAGCGTCTT	581					
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654)		
QY	AG158197	Pan troglodytes DNA, clone: RP43-023F15.T7, genomic survey	AG158197	GI:1668785	GSS.	Pan troglodytes male lymphocytes DNA, clone: RP43-023F15.T7.	Male BAC library clone: RP43-023F15.T7.	Pan troglodytes	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Tokoki, Y., Watanabe, H., and Sakaki, Y.	BAC end sequences of library RPCI-43 unpublished	2 (bases 1 to 654		

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	22	TTCCGGGACACACCATAGCGTGGAGAGACCCCTCTACGTACGTGGTAAGACAAAGGCCATAAAT	81													
OY	731	ATCTTGTCTTTGAAGCCGATCTCTGGACACACAGATGGTTATCAAGGCGATCAGTCTTCA	790													
Db	82	ACCTTGTATTCGAAGCCACACGGGAGACAGAAAAAGGATATACAGGGGAAGAACTTTAT	141													
OY	791	ATATATAAAGCTTACTATGCGGGAAGTGACGTCCTTCCGATGAAAGAAAAATTAACCTGC	850													
Db	142	TTTAACAAGGCTTACTACGCGCGGCGCACAGAACTTCTCCGTAAGAAAGCCAGAAAGCTTC	201													
OY	851	TTCAAGTCTCTTAAAAAACAAATATCTTCTTTAGGGAATGTGATATAGGCAATTTCTGAAT	910													
Db	202	AGCAGAGGCGCTTAAAAAACGGCGATCTGAGTTAGGAAGCGCGCCCTGGTATCTATAGAGT	261													
OY	911	TGCGCGATGACTATACAGTGAAGTGTATTAAGAACCTTAGTCGCATCAACACAGTAG	970													
Db	262	TAAATTAATGATACATTCATGAAAAAGTAAAGAACGCGCTGATCATCTTCAACACGGTAA	321													
OY	971	CAGATGAAGTGGAAAGCGGCGCAATATATTTAAATGAATAAATAAATGATATCTACCG	1030													
Db	322	CTGATGAATATGAGAGCGCGGAGATGTTTCAAAATGAAGGCCAAATAGTACTTGTTCATG	381													
OY	1031	ATTCAAGAGATATCCAAATATGACGATGTATGGAATTAAGACAAAGATTTTATATGCTAG	1090													
Db	382	ATTCAAGCGGGTTCAAAATATGACGATGATGATTTAATCAACAGATATTTACATGCTTG	441													
OY	1091	GGCCCGGAGCGACTCTTAAATGAGCCGCACAAACCCGATAAATGAAGTGAAGCTTGAT	1150													
Db	442	GTTATGTATCAACTCTTTAACCGGCGCTTCAACAGCGCTGAAACAAACAGCGGCTGTGC	501													
OY	1151	TGAACATGAATCTTGACCCCTGCTGATCTCACACACACTTACTCTCATTTGGCGTATCCGC	1210													
Db	502	TGCAATATGGCTTGATATCAACAGATGACATTCATCTACTCTCATCTTCCAGTGGCGC	561													
OY	1211	ACCCGTGAAGTAAATATGCTGCTACTACACAAATTAATGACAAATGAAGGCTTCT	1264													
Db	562	AAGCCAAAGCAACATGTGGTTTTCACACACTCATGACAAACAGAGGCTTCT	615													
RESULT 4	AG145469	AG145469 687 bp DNA linear GSS 08-JAN-2002														
LOCUS		Pan troglodytes DNA, clone: RP43-006M11.T7, genomic survey														
DEFINITION		Sequence.														
ACCESSION		AG145469														
VERSION		AG145469.1 GI:16675147														
KEYWORDS		GSS.														
SOURCE		Pan troglodytes male lymphocytes DNA, clone: lib:RP43-006M11.T7.														
ORGANISM		Male BAC library clone:RP43-006M11.T7.														
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;														
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.														
TITLE		1														
JOURNAL		Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,														
REFERENCE		2 (bases 1 to 687)														

[illegible]

OY		451	GGCCACAGTCCTTATTCTGTATACGAGATTACTCAGGTAATCCTGAAGAATGGTGAAACCGGT	510
Db		122	GAAAAAATCGTTATTCTTACACGACTATTCGTTTTAACAFPA-----	165
OY		511	GCTGGTAAACCATAATCATTTCAACTGCCTCACTAACTAACTATGCCAGCCGATGACATACA	570
Db		166	--CGGCAACAAAGCGCTGCACACAGCGCAGGTAATGTGTCAAATCTGAT---GACACA	220
OY		571	CTTAAAGTCGATGGAGTATCTGATCATTAATCTGTCTTTGATGGCGGAGACGGTACAGTT	630
Db		221	CTCAAAAATCAACGGAGTGTGGAGATCACAAAACGATTTTGTGA---CGGAGACGGAACACA	277
OY		631	TATCAAAATATTCCAGCATTTATGATGATGAAGGCAAGTGGATTCCAGTGATATACCATCT	690
Db		278	TATCAGAAAGTTTCAGCAGTTTATGATGATGAAGGCAATTTATACATCTCGGACACACCATACG	337
OY		691	TTAAGAGACCCCTCACTATGTGAAGATPAAAGGGCCATAATATCTTGTCTTTGAAGCGAAT	750
Db		338	CTGAGAGACCCCTCACTACGTGTAAGAGACACAGCCATTAATACCTGTATTTGGAAGCCAC	397
OY		751	ACTGGAACACAGATGTTATTCAGAGCGATCAGTCTTTCAATATATTAAGCTTACTANTGC	810
Db		398	ACGGGAACAGAAAACGATACCAAGGCGAMAATCTTTATTACAAAGCCCTACTACAGCG	457
OY		811	GGAGTGAAGCTCTCTCTTCGCAATGAATAAATAATTAACGCTTCAAAAGCTCAAAAAACAA	870
Db		458	GGCGGCACGAACTCTCTTCGTAAGAACCCAGAAAGCTTTCAGCAAGCGCTCAAAAAAGCG	517
OY		871	ATTGCTCTTAGCGAATGTGCATATTAGCAATTGTGTAATTGGCCGATGACTATACAGTG	930
Db		518	GATGCTGATGTTAGGGAACGGGCGCCCTGGTATCAGATTAAGTTAATATGATTACCAATTG	577
OY		931	AAAAGTGTATTGAACACATTAGTGCATCAACACAGTAGACGATGAGAGTGAAGCGGCC	990
Db		578	AAAAAGATTAAGAAAGCCGCTGATCATCTTATATACAGGTAAGTGAATGAGAGCGCGC	637
OY		991	AATATATTTAATAATGA 1006	
Db		638	AATGTTTTCAAAATGA 653	
RESULT 7				
LOCUS	AG165661	688 bp	DNA linear	GSS 09-JAN-2002
DEFINITION	Pan troglodytes DNA, clone: RP43-033D03.T7, genomic survey sequence.			
ACCESSION	AG165661			
VERSION	AG165661.1	GI:16695339		
KEYWORDS	SOURCE			
ORGANISM	Pan troglodytes male lymphocytes DNA, clone_1lb:RCI-43 Chimpanzee Pan troglodytes			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.			
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.			
TITLE	BAC end sequences of library RPCI-43			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 688)			
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-Chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://npg.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)			
COMMENT	Clones are derived from the chimpanzee BAC library RPCI-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.			
PRIMERS				

Sequencing: T7	
LIBRARY	Vector : pBACE3.6
FEATURES	R.Site 1 : EcoRI
source	R.Site 2 : EcoRI.
	Location/Qualifiers
	1..688
	/organism="Pan troglodytes"
	/db_xref="taxon:9598"
	/clone="RP43-033D03.T7"
	/sex="male"
	/cell_type="lymphocytes"
	/clone_lib="RPC1-43 Chimpanzee Male BAC library"
BASE COUNT	227 a 160 c 154 g 147 t
ORIGIN	
Query Match	21.0%; Score 286.6; DB 17; Length 688;
Best Local Similarity	68.4%; Pred. No. 2.4e-65;
Matches	451; Conservative 0; Mismatches 184; Indels 24; Gaps 3
Oy	171 AATAGATTAGATGATGAGATGAGTACGTGGCCACTGCAAAAGCGTGATGGCGGCAAA 230
Db	53 AAAAGACTTGTATGTGTGGACAGCTGGCCGCTGCAAAACGTGACGGACAGTACGCTGA 112
Oy	231 TTATCAGTATATCATCATGCTGCTCCGCTTTAGCAGGTGACCCAAAACAGTATGATAC 290
Db	113 ATACACACGGCTATCCGTTGTGTGTTGCTTCCGGGAAAGCCGGAAGACGCTGATGACAC 172
Oy	291 TCCACTTATTTATCTATCAAAAAGTGGGTGATCATCGATTGACAGCTGGAAAATGC 350
Db	173 ATCAATCTACATGTTTATTAACAAGGTGGGACACTCACTCAACAGCTGGAAGAACGC 232
Oy	351 TGAAGAGTATTTGAAGATATGATGATAAATTTGTTCCAATGATCCGTAATCTTAATATCA 410
Db	233 GGGCCGTGCTTTAAAGACAGAGATAGTGCAGCCCAACAGTCCGATCTGAAGATCA 292
Oy	411 AACACAGAGTGTGTAGGTTCTGCTACTTTAAACAAAGATGSCCAAGTCCGTTATTCTCA 470
Db	293 GACGCMAAGAAATGGTCCGGTCTGCACACCTTTACATCATCAGGAAAATCCGTTATTCTCA 352
Oy	471 TACAGATTACTCAGTAACTCTGGAAGATGTGTGAACCGGTGCTGTATCAACAATATTTTC 530
Db	353 CACTGACTATTCGGTAAACATTA-----CGCAACAAAAGCCTGAC 394
Oy	531 AACTGCTCAAGTAAACTTATCCACCGGAGATGACAGTACACTTAAAGTGATGAGATTC 590
Db	395 AACAGCGCAGGTAATGTGTCAAAATCTGAT---GACACACTCAAAATCAACGAGGTGA 451
Oy	591 TGATCATTAATCTGCTTGTGATGGGGAGAGCGGTACAGTTATCAAAATATGACCAATT 650
Db	452 AGATCACAAAACGATTTTGA---CGGAGACGGAAGAAAACATATCAAGACGTTACACAGTT 508
Oy	651 TATCGATGAAGGCAAGTGGATTTACAGTATGATTAACCATCTTTAAGAACCCCTCACTATGT 710
Db	509 TATCGATGAAGCAATTTATACATCCGGGACACACCATACCGTGAAGACCCCTCACTAGCT 568
Oy	711 TGAAGATGAAGGGCCATAATATCTGTCTTTGAACGGAATACTGGAACAACAGATGTTA 770
Db	569 TGAAGACAAAAGCCATAATACCTTTGTATTCGAACCCCAACAGGGAAGAAAAGGATTA 628
Oy	771 TCAAGGCGATAGCTCTTCATTAATTAAGCTTACATGAGCGGAAGTGAACGTTCTTCC 829
Db	629 CCAAGGCGAAGATCTTTATTTTACAAAGCGTACTAGCGGCGGACAGAACTTCTTC 687
RESULT 8	
AG158505	656 bp DNA linear GSS 09-JAN-2002
LOCUS	Pan troglodytes DNA, clone: RP43-023M01.T7, genomic survey
DEFINITION	
ACCESSION	AG158505
VERSION	AG158505.1
KEYWORDS	GSS: 1668183

SOURCE	Pan troglodytes male lymphocytes DNA, clone_1lb:RPI-43 Chimpanzee Male BAC Library clone:Rp43-023M01.T7.
ORGANISM	Pan troglodytes
REFERENCE AUTHORS	Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
TITLE	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y. BAC end sequences of library RPI-43 Unpublished 2 (bases 1 to 656) Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y. Direct Submission Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan 1-7-22 Suehitto-Chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@sc.riken.go.jp, URL:http://npg.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT	Clones are derived from the chimpanzee BAC library RPI-43 This BAC end was generated during the Rad process and may have higher chance of clone tracking errors.
PRIMERS	
LIBRARY	Sequencing: T7
Vectors : PBACE3.6	
R.Site 1 : EcoRI	
R.Site 2 : EcoRI	
Location/Qualifiers	
/organism="Pan troglodytes"	
/db_xref="taxon:9598"	
/clone="RP43-023M01.T7"	
/sex="male"	
/cell_type="lymphocytes"	
/clone_lib="RPI-43 Chimpanzee Male BAC Library"	
BASE COUNT	229 a 144 c 140 g 143 t
ORIGIN	
Query Match	20.8%; Score 284.4; DB 17; Length 656;
Best Local Similarity	69.6%; Pred. No. 9,2e+65;
Matches 416:	Conservative 0; Mismatches 176; Indels 6; Gaps 2
OY	491 CTGAAGATGTGGAAACCGCTCCTGTTAACCAATCATTTCAACTGCCTCAAGTAACCTTAT 550
D.	65 CTGACTATTTCGGTGAAACATTATACGCCAACAAGCCATCACAAAGCCGAGGTAATGTGT 124
OY	551 CCCAGCCGATGACAGCTACTCTTAAGTCGATGAGTCTGCATCATTAATCTGCTCTTG 610
Dd	125 CAAATCTGAT--GACACACTCAAATACTCAACGGAGTGGAAGATCACAAAAGATTTTG 181
OY	611 ATGGCGAGACGCTACAGTTTATCAAAATATTTCAGCAATTTATCGATGAGGCAAGTGA 670
Dd	182 A---CGGAGACGAAAAACATATGACAGCTTACAGCAGTTTATGATGAAGCAATTTATA 238
OY	671 TTTCAGCTGATTAACCATTACTTTAAGAAGCCCTCACTATGTTGAAGATAAGGCCCTAAT 730
OY	731 ATTCTGCTTTGAAGCGAATCTCTGAAACAACAGATGTTATCAAGGCGCATGACTTTCCA 790
Dd	299 ACCTTGTTTTCGAAAGCCAAACAGGGAACAAGAAAAGGATACCAAGGCGCAAGATCTTAT 358
OY	791 ATAATTAAGCTTACTATGGCGGAAGTAGCAGCTCTTCTCCAAATGAAAAAAATAAATC 850
Dd	359 TTAAACAAAGCTACTACGGCGCGGCAACGAATCTTCTCGTAAAGAAAAGCCAGAACTTC 418
OY	851 TTCGAAGCCCTAAAAAAACAATTCCTCTTTAAGGAAGTGTCGATTTAGCGATTTGGTGAAT 910
Dd	419 AGCAGAGCGCTAAAAAAGCGCATGTCTAGTTTACGAGAGGGGCCCTCGTATCATAGAGT 478
OY	911 TGCGCGATGACTATACACTGAAGAAGTGTTATGAACAACATTGTGCGATCAACAACAGTAG 970

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	FEATURES	ORIGIN	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0:	Mismatches	191:	Indels	25:	Gaps	4:
Db	479	TAAATTAATGATTACACATTAAGAAAAAGTAATGAAGCCGCTATCATCTTCAACACGGTAA	538																									
Oy	971	CAGATGAAGTGGAAAGCCGCAATATATTTAAATGAATTAATTAATGTATCTATTCACGG	1030																									
Db	539	CTGATTAATGAAATCGCGCGCAATGTTTTCATTAAGAAAGCAAGCAATGTACTTGTCACTG	598																									
Oy	1031	ATTCAGAGGATTCACAAAATGACGAGTGAATTAACGACAAAGATCTTTATATGCT	1088																									
Db	599	ATTCACGCGGTTCAAAAATGACGATCATGTGATTAACCTCAACAGATATTTATACATGCT	656																									
RESULT 9	LOCUS	AG019951/c	678 bp	DNA	linear	GSS 16-OCT-1999																						
DEFINITION	Homo sapiens	genomic DNA, 21q region, clone: B293N7 H058(RP),																										
ACCESSION	AG019951	genomic survey sequence.																										
VERSION	AG019951.1	GI:6045895																										
KEYWORDS	GSS.																											
SOURCE	Homo sapiens	DNA, clone: B293N7 H058(RP).																										
ORGANISM	Homo sapiens																											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																											
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.																											
TITLE	Hattori, M., Fujiyama, A., Ishii, K., Toyoda, A., Taylor, T.,																											
JOURNAL	Park, H.-S., Yada, T., Wetanabe, H. and Sakaki, Y.																											
REFERENCE	Homo sapiens genomic DNA, chromosome 21q																											
AUTHORS	Published Only in Database (1999)																											
TITLE	2 (bases 1 to 678)																											
JOURNAL	Hattori, M., Fujiyama, A., Ishii, K., Toyoda, A., Taylor, T.,																											
REFERENCE	Park, H.-S., Yada, T., Wetanabe, H. and Sakaki, Y.																											
AUTHORS	Direct Submission																											
TITLE	Submitted (16-OCT-1999) Masahira Hattori, RIKEN Genomic Sciences																											
JOURNAL	Center(GSC) c/o Kitasato University, 1-15-1 Kitasato, Sagamihara																											
REFERENCE	228-8555, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp,																											
AUTHORS	Tel:042-778-9923, Fax:042-778-9924)																											
FEATURES	Location/Qualifiers																											

Oy	530	AAACTGCTCAAGTAACCTTATCCAGCGCGGTGACGTACACTAAAGTCGATGACAT	589
Db	314	CAACAGCGCAGGTAAATGCTCAAAATCTGAT---GACACATCAAAATACAGCACTGG	258
Oy	590	CTGATCATTAATCTGCTCTTGATGGCGGAGACGGTACAGTTTATCAAAATATTCAGCAAT	649
Db	257	AAGATCACAAAACGATTTTTTGA---CGAGACGGAAAAACATATACAGGTTCCAGAGT	201
Oy	650	TTATTCGATGAAGGCAAGTGGATTTTCAGGTGATTAACCATACTTTAAGACACCTCAGTATG	709
Oy	710	TTGAAGATTAAGGGCCATTAATATCTTGCTTTGAAGGAAATACGAAACAACAGATGGTT	769
Db	140	TTGAAGACAAGGCCATTAATATCTTGATTCGAAGCCAACAGGGAACGAAAACGGAT	81
Oy	770	ATCAAGCGCATCTCTTCAATATAAAGCTTACTATGCGGAAGACAGCTCTTCCTTCC	829
Db	80	ACCAAGTGGGAAGATCTTTATTTTAAACAAGGCTACTACGGGGCGGACAGCACTTCTCC	21
Oy	830	AGATGAAAAAATAAATTAACCT	848
Db	20	GTAAGAAGAACCCAGAGCT	2
RESULT 10			
LOCUS	AG001039/c		
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: S39BG29, genomic survey sequence.		
ACCESSION	AG001039		
VERSION	AG001039.1		
KEYWORDS	GSS.		
SOURCE	Homo sapiens DNA, clone: S39BG29.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.		
TITLE	Homo sapiens genomic DNA, chromosome 21q		
JOURNAL	Published Only in Database (1997)		
AUTHORS	Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-OCT-1997) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)		
FEATURES			
SOURCE	Location/Qualifiers		
	1..711		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="21"		
	/map="21q"		
	/clone="S39BG29"		
BASE COUNT	145 a 158 c 166 g 241 t	1	others
ORIGIN			
Query Match	19.9%; Score 271; DB 17; Length 711;		
Best Local Similarity	67.3%; Pred. No. 3.5e-61;		
Matches	466; Conservative % 0; Mismatches 201; Indels 25; Gaps 5;		
Oy	328	TCGATTACACAGCTGGAAAAATGCTGCAAGACTATTTAAAGATATGCAATAATTTGTTCCA	387
Db	708	TCATTCATTCACAGCTGGAAAAAGCGGCGCGTGTCTTTAAACACGCGATCTGCACCC	649
Oy	388	AATGATCCGATCTTAAATATCAACACAGAGTGTGACAGTTCTGCCTTAAACCAA	447
Db	648	AACGATCCG-ATCTGGAAGATCAGACCA-GAATGTCGCGTTCTGCAACCTTTACATCT	591
Oy	448	GATGCCCAAGTCCGTTTATTTCTATACAGATTACTACAGTAATCCTGAAAGATGCTGAAC	507


```

Db      590 GACGGAATAATCCGTTTATTTCTACACTGACTATTCGCGTAACAT----- 546
QY      508 GGTGCTGTAAACCAATCATTTTCAACTGCTCAAGTAAATTAATCCAGCCGAGTGCAGT 567
Db      545 --TACGGCCAAACAAAGCCCTGACACAGCCAGGTAATATGTCAAAATCTGAT---GAC 491
QY      568 ACACTTAAAGTCATGAGTATCTGATCAATTAATCTGCTCTTGTATGAGGAGACGGTACA 627
Db      490 ACACCTCAAAATCAACGAGTGAAGATCACAAAACGATTTTGA---CGGAGACGGAAAA 434
QY      628 GTTATCAAAATATTCGCAATTATCGAATGAGGAGGAGGATTTAGGTGATACCAT 687
Db      433 ACATATGAGAACGTTTCAGCAGTTATGATGAGGAAATTAATACATCCGGCAGCAACCAT 374
QY      688 ACTTTAGAGACCCCTCAGTATGTTGAAGATAGGCGCATTAATCTTGTCTTGAAGCG 747
Db      373 ACGGTGAGAGACCCCTCAGTATGTTGAAGATAGGCGCATTAATCTTGTATTCGAGGCC 314
QY      748 AATACCTGGAACAAAGATGTTATCAAGCGGATCAGTCTTTCATTAATAAGCTTACTAT 807
Db      313 AACACGGGAACAGAAAACGGATACACAGCGGAAATCTTATTTAACAAGCGTACTAC 254
QY      808 GGGGAGATGACGCTCTTCCAGAAATGAAAAAATAACTGCTTCAAGTCCATAAAA 867
Db      253 GGGGCGGCGACGAACTTCTCCGTAAGAAAGCCAGAGCTTCAGCAGCGCTTAAAAA 194
QY      868 CAAATTCCTCTTTCAGCAATGTTGATGAGCATTTGATTTGATGCGCATGACTATACA 927
Db      193 CGGATGCTGAGTATGAGGAGCGGCGGCGGATATCATAGATTAAATATGATTTACACA 134
QY      928 GTGAAAGCTTATGAAACCATTAATGTCATCAACACAGTACGAGTGAAGTGAACGC 987
Db      133 TTGAAAAAGTAAATGAAGCGCTGATCATTCAACACGCTTAACCTGATGAATCGAGCG 74
QY      988 GCCAATATTTTAAATGAATTAATTAATGTA 1019
Db      73 GCCAATTTTTCAAAATGAAGCGCAATGTA 42

RESULT 11
AG156831
LOCUS   Pan troglodytes DNA, clone: RP43-021J04.T7, genomic survey
DEFINITION
ACCESSION AG156831.1 GI:16686509
VERSION    AG156831.1
KEYWORDS   Pan troglodytes male lymphocytes DNA, clone:lib:RP43-43 Chimpanzee
SOURCE     Male BAC Library clone:RP43-021J04.T7.
ORGANISM   Pan troglodytes
REFERENCE  1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of library RP43-43
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 669)
TITLE      Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
           and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
           1-7-22 Suehiro-cho,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
           (E-mail:chimbesc@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
           tel:81-45-503-9111, fax:81-45-503-9170)
COMMENT    Clones are derived from the chimpanzee BAC library RP43-43 This BAC
           end was generated during the R&D process and may have higher chance
           of clone tracking errors.
PRIMERS
           Sequencing: T7
LIBRARY   Vector
           : pBACe3.6

```

```

FEATURES
Source
R.Site 1 : EcorI
R.Site 2 : EcorI.
Location/Qualifiers
1..669
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-021J04.T7"
/sex="male"
/cell_type="lymphocytes"
/clone.lib="RP43-43 Chimpanzee Male BAC Library"
BASE COUNT 227 a 153 c 143 g 146 t
ORIGIN
Query Match 19.6%; Score 267; DB 17; Length 669;
Best Local Similarity 65.5%; Pred. No. 4e-60;
Matches 390; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY      749 ATACTGGAACACAGATGTTATCAAGCGCATGCTTTCAATTAATAAGCTTACTATG 808
Db      34 ACAGGGGAAACGAAAACGGATACCAAGGCGAAGAACTTTATTTAACAAGCGTACTACG 93
QY      809 GCGGAGTGAAGTCTTCTTCCAGATGAAAAAATAACTGCTTCAAGTCCATAAAAC 868
Db      94 GCGGCGGACGAACTTCTTCCGTAAGAAAGCCAGAAAGCTTCAGCAGCGCTTAAAAAC 153
QY      869 AATTCCTCTTTCAGCAATGTTGATGAGCATTTGTAATGCGCGATGACTATACAG 928
Db      154 GCGATGCTGATGATGAGGAGCGGCGGCGGATATCATAGTTAAATATGATTTACAT 213
QY      929 TGAAGAAGTATTAAGAACCTTAATGTCATCAACACAGTACGATGATGAAGTGAACGG 988
Db      214 TGAAGAAGTAAATGAAGCGCGTGTATCACTCAACACGCTTAATGATGATGATGACGCG 273
QY      989 CCAATATTTTAAATGAATTAATTAATGATATTCACGATTCACAGAGATCCAAAA 1048
Db      274 CGATGTTTCCAAATGAAGGCAATGACTTGTTCATGCTGATTCACGCGTTCAAAA 333
QY      1049 TGACAGTGAATGAATTAACGACAAAGATTTTATGCTAGAGCCCGGAGCGACTCT 1108
Db      334 TGACGATGATGATGATTAATTAACGACGATTTTATGATGATGATGATGATGATGAT 393
QY      1109 TAAATGCGCCACCAACCCGATTAATGAACAGTGTGATTTGAACATGATCTTGACC 1168
Db      394 TAAACGCGCTTCAAGCGCTGTGAACAAAACAGCGCTGTGCTGAGTAAGAGCTTGATC 453
QY      1169 CTGCTGATTCACACACACTTACTCTCATTTGGCGTATCCGCGACCTGGAAGTAATAG 1228
Db      454 CAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 513
QY      1229 TGTACTACACAGTATATGACGATAGAGGCTTTCTATCCAGAACATCACTCTACCTGC 1288
Db      514 TGTATTACCAACACTCATGATGACAAACAGAGGCTTTCTGAGATTAAGAGCAACATTTG 573
QY      1289 GGCACAGCTGGGTTAATATTAAGGCTGACACATCTGAGAGAGAAATAG 1343
Db      574 GCGCAAGCTTCTTAATGATGATCAAGGCAATTAAGCAATCCGTTGCAAAAACAG 628

RESULT 12
AG0791204
LOCUS   Human genomic clone Plate-4555 Col-19 Row-B, DNA sequence.
DEFINITION
ACCESSION AG0791204
VERSION    AG0791204.1 GI:5698916
KEYWORDS   GSS.
SOURCE     Homo sapiens
ORGANISM   human.
REFERENCE  1
AUTHORS    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
           Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

```


QY 1252 AATAGAGCTTCTATCCAGAACATCATCTCAGCTGCGGAGAACGTTGGGTTATATTT 1311
 DB 201 AACAGAGCTTCTTCCAGGATTAAGAACATTTTCCCAAGCTTTTAAATGAACATC 142
 QY 1312 AAAGGCTCGACATCTGAGAGAGAAATAG 1343
 DB 141 AAAGCAATTAACATCGTGTCTCAAAACAG 110
 RESULT 14
 AL575209 934 bp mRNA linear EST 16-FEB-2001
 LOCUS AL575209 LTI.NFL006.PL2 Homo sapiens cDNA clone CSOD1062YP02 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL575209
 VERSION AL575209.1 GI:12936153
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 934)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1. 934
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSOD1062YP02"
 /clone_lib="LTI.NFL006.PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang life technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 298 a 196 c 223 g 209 t 8 others
 ORIGIN
 Query Match 18.3%; Score 250; DB 9; Length 934;
 Best Local Similarity 64.8%; Pred. No. 1.5e-55;
 Matches 367; Conservative 2; Mismatches 197; Indels 0; Gaps 0;

QY 1078 GTTATATTCAGAGCCCGAGCGATCTCTTAATGAGCCACACACCGATTAATGAA 1137
 DB 323 ATTATATTCAGTGTGTTATGATCAACATCTTTAAACGGCCCTTACAAAAMAAAAA 382
 QY 1138 ACTGACCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
 DB 383 AAAGGCTGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
 QY 1198 TCGGATATCCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
 DB 443 TTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
 QY 1258 GCGTCTATCCAGAACATCATCTCAGCTGCGGAGAACGTTGGGTTAATTAAGG 1317
 DB 503 GCGTCTTCTGAGATTAAGAACATTTGGCGCAAGCTTTAATGACATCAAGAGC 562
 QY 1318 TCTGACATCTGAGAGAGAAATAG 1343
 DB 563 AATAAACAATCCGTTGTCAAAACAG 588
 RESULT 15
 AG185716 696 bp DNA linear GSS 09-JAN-2002
 LOCUS AG185716
 DEFINITION Pan troglodytes DNA, clone: RP43-059006.T7, genomic survey
 sequence.
 ACCESSION AG185716
 VERSION AG185716.1 GI:16715396
 KEYWORDS GSS.
 SOURCE Pan troglodytes male lymphocytes DNA, clone_lib:RP43-Chimpanzee
 Male BAC library clone:RP43-059006.T7.
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
 REFERENCE 1
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Tokoi, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of library RP43-43
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 696)
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Tokoi, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chiimpes@sc.riken.go.jp, URL: http://hnp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 COMMENT Clones are derived from the chimpanzee BAC library RP43-43 This BAC
 end was generated during the R&D process and may have higher chance
 of clone tracking errors.
 PRIMERS
 Sequencing: T7
 LIBRARY
 Vector : pBACE3.6
 R.site 1 : EcoRI
 R.site 2 : EcoRI
 Location/Qualifiers
 1. 696
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="RP43-059006.T7"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RP43-Chimpanzee Male BAC Library"
 BASE COUNT 237 a 162 c 148 g 146 t 3 others
 ORIGIN
 Query Match 18.2%; Score 248.6; DB 17; Length 696;
 Best Local Similarity 64.6%; Pred. No. 3.2e-55;
 Matches 459; Conservative 0; Mismatches 200; Indels 51; Gaps 4;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 07:25:42 ; Search time 206.939 Seconds
(without alignments)
9679.375 Million cell updates/sec

Title: US-09-986-682B-4
Perfect score: 1365
Sequence: 1 ATGACAGCGCGGACATACAA.....CCGACAGACATTCCTCA 1365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1365	100.0	1365	10	US-09-986-682B-4
2	1365	100.0	2408	10	US-09-986-682B-5
3	617.2	45.2	1449	10	US-09-974-300-2644
4	578.6	42.4	2253	10	US-09-993-292A-19
5	578.6	42.4	8908	10	US-09-993-292A-18
6	577	42.3	4938	9	US-10-055-794-1
7	577	42.3	6525	9	US-10-055-794-3
8	577	42.3	10317	10	US-09-849-866-1
9	135.6	9.9	368	10	US-09-864-761-21746
10	73.2	5.4	376	10	US-09-864-761-5018
11	72.4	5.3	1146	10	US-09-778-844-136
12	61.2	4.5	4634	10	US-09-995-587A-10
13	52	3.8	2367	10	US-09-995-587A-2
14	52	3.8	2394	10	US-09-995-587A-3
15	52	3.8	2592	10	US-09-995-587A-4
16	44	3.2	58985	9	US-09-901-152-3
17	41.8	3.1	5301	7	US-08-781-986A-443
18	38.6	2.8	921	10	US-09-815-242-4568
19	38.6	2.8	957	10	US-09-815-242-8609

C 20	38.4	2.8	5086	9	US-09-764-891-9794	Sequence 9794, Ap
21	38	2.8	314	10	US-09-833-381-190	Sequence 190, App
22	38	2.8	548	10	US-09-874-300-3392	Sequence 3392, Ap
23	37.8	2.8	3831	12	US-10-051-952-5	Sequence 5, Appl
24	37.4	2.7	8761	9	US-10-098-841-11	Sequence 11, Appl
25	37.4	2.7	10758	12	US-10-044-090-61	Sequence 61, Appl
26	37.2	2.7	3996	6	US-10-087-464-42	Sequence 42, Appl
27	37	2.7	1799	7	US-08-781-986A-447	Sequence 447, App
28	36.6	2.7	7101	9	US-10-011-366-9	Sequence 9, Appl
29	36	2.6	1795	10	US-09-764-864-145	Sequence 145, App
30	36	2.6	2971	10	US-09-815-242-3842	Sequence 3842, Ap
31	36	2.6	5136	10	US-09-070-927A-270	Sequence 270, App
32	35.6	2.6	556	10	US-09-864-761-8619	Sequence 8619, Ap
33	35.4	2.6	439	10	US-09-864-761-20174	Sequence 20174, A
34	35.4	2.6	3317	9	US-10-055-412B-1	Sequence 1, Appl
35	35.4	2.6	378361	9	US-09-901-136-3	Sequence 3, Appl
36	35	2.6	327	10	US-09-864-761-28059	Sequence 28059, A
37	35	2.6	456	10	US-09-864-761-11468	Sequence 1, Appl
38	35	2.6	2058	9	US-10-216-289-3	Sequence 3, Appl
39	35	2.6	14483	9	US-10-216-289-3	Sequence 3, Appl
40	35	2.6	82938	9	US-09-818-657-3	Sequence 3, Appl
41	34.8	2.5	470	10	US-09-864-761-2617	Sequence 2617, Ap
42	34.6	2.5	513	9	US-10-123-155-192	Sequence 192, App
43	34.6	2.5	772	9	US-09-938-842A-4312	Sequence 4312, Ap
44	34.6	2.5	8996	9	US-10-239-676-212	Sequence 212, App
45	34.6	2.5	12313	9	US-10-171-311-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-986-682B-4
Sequence 4, Application US/09986682B
Patent No. US20020115182A1
GENERAL INFORMATION:
APPLICANT: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU
KENKYUJO
TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOPURANOSIDASE
ACTIVITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/986, 682B
FILING DATE: 06-NO. US20020115182A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 170,630/1996
FILING DATE: 10-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-737-3528
TELEFAX: 202-737-3528
INFORMATION FOR SEQ. ID NO. 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1365 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
us-09-986-682b-4

Query Match 100.0%; Score 1365; DB 10; Length 1365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 ATGAACAGCGGGGACTACAGAGACATATGTTTGGCCATATTATACAGCGCTGACATG 60
Db 1 ATGAACAGCGGGGACTACAGAGACATATGTTTGGCCATATTATACAGCGCTGACATG 60
Oy 61 CTAAAAATTCAGAGACAACAACAGTCTCAATTTAAAGTCCCTCAATTCATCATCA 120
Db 61 CTAAAAATTCAGAGACAACAACAGTCTCAATTTAAAGTCCCTCAATTCATCATCA 120
Oy 121 GCATCAAAAACATTTGATGGGCAAAAGGATATGTAAGTACAGCAACTTAAATGATTTA 180
Db 121 GCATCAAAAACATTTGATGGGCAAAAGGATATGTAAGTACAGCAACTTAAATGATTTA 180
Oy 181 GATGATATGGATAGCTGGCCACTGCAAAACGCTGATGTCGGCAAAATTCATGTA 240
Db 181 GATGATATGGATAGCTGGCCACTGCAAAACGCTGATGTCGGCAAAATTCATGTA 240
Oy 241 TATCAATCGTCTCGCTTTAGACAGGTGACCCAAAAACAGTATGATGATCTCCATCAT 300
Db 241 TATCAATCGTCTCGCTTTAGACAGGTGACCCAAAAACAGTATGATGATCTCCATCAT 300
Oy 301 TTATTTCTCAAAAAGGCGGATACATCATGATGACAGTGGAAAAATGCTGAAAGATA 360
Db 301 TTATTTCTCAAAAAGGCGGATACATCATGATGACAGTGGAAAAATGCTGAAAGATA 360
Oy 361 TTGAGAGATGATGATTAATTTGTTCCAAATGATCGTATCTTAAATATCAACACAGAG 420
Db 361 TTGAGAGATGATGATTAATTTGTTCCAAATGATCGTATCTTAAATATCAACACAGAG 420
Oy 421 TGGTCAAGTTTCTGCTACTTTAACCAAGATGGCCAGTCCGTTATTTCAACAGATTAC 480
Db 421 TGGTCAAGTTTCTGCTACTTTAACCAAGATGGCCAGTCCGTTATTTCAACAGATTAC 480
Oy 481 TCAGTATTCCTGGAAGATGTGAACCGGTGCTGTAACCAATCATTTCAACAGCTCAA 540
Db 481 TCAGTATTCCTGGAAGATGTGAACCGGTGCTGTAACCAATCATTTCAACAGCTCAA 540
Oy 541 GTAACCTTATCCAGCCGGATGACGATCACTTAAAGTCAGTATGATCATATAA 600
Db 541 GTAACCTTATCCAGCCGGATGACGATCACTTAAAGTCAGTATGATCATATAA 600
Oy 601 TCTGCTTTGATGGCGGAGAGCGTACATTTCAAAATATTCAGCAATTTATCGATGAA 660
Db 601 TCTGCTTTGATGGCGGAGAGCGTACATTTCAAAATATTCAGCAATTTATCGATGAA 660
Oy 661 GGAAGTGTGATTTAGGATGATACCATTTTAAAGAGCCCTCACTATTTGAAGATTAAG 720
Db 661 GGAAGTGTGATTTAGGATGATACCATTTTAAAGAGCCCTCACTATTTGAAGATTAAG 720
Oy 721 GGGCATAAATATCTGCTTTGAAGCGAATATCGAACACAGATGTTATCAAGCGCAT 780
Db 721 GGGCATAAATATCTGCTTTGAAGCGAATATCGAACACAGATGTTATCAAGCGCAT 780
Oy 781 CAGCTTTTCAATATTAAGCTTACTATGCGGAAGTGACGCTCTTCCAAATGAAGAAA 840
Db 781 CAGCTTTTCAATATTAAGCTTACTATGCGGAAGTGACGCTCTTCCAAATGAAGAAA 840
Oy 841 AATAAAGTCTTCAAAAGCTTAAAAAACAATTTGTTTAAAGGAATGGTCAATTAGC 900
Db 841 AATAAAGTCTTCAAAAGCTTAAAAAACAATTTGTTTAAAGGAATGGTCAATTAGC 900
Oy 901 ATTGTTGAATGGCGGATGATATACAGTGAAGGTGTTATGAACAACATTTAGTGCATCA 960
Db 901 ATTGTTGAATGGCGGATGATATACAGTGAAGGTGTTATGAACAACATTTAGTGCATCA 960
Oy 961 AACACAGTACGATGAAGTGAACGCCCAATATTTAAATGAATTAATTAATGCTAT 1020
Db 961 AACACAGTACGATGAAGTGAACGCCCAATATTTAAATGAATTAATTAATGCTAT 1020

```

```

Db 961 AACACAGTACGATGAAGTGAACGCCCAATATTTAAATGAATTAATTAATGCTAT 1020
Oy 1021 CTATTACAGGATTTACAGAGATCCAAAATGACGATGATGATTAACGACAAAGATGT 1080
Db 1021 CTATTACAGGATTTACAGAGATCCAAAATGACGATGATGATTAACGACAAAGATGT 1080
Oy 1081 TATATGCTAGGGGCGGAGCGGACTCTTAAATGCGCCACACAAACCGATTAATGAAC 1140
Db 1081 TATATGCTAGGGGCGGAGCGGACTCTTAAATGCGCCACACAAACCGATTAATGAAC 1140
Oy 1141 GCACTTATTAAGAACATGATCTTGACCCCTGCTGATCTCACACACATTTACTCATTC 1200
Db 1141 GCACTTATTAAGAACATGATCTTGACCCCTGCTGATCTCACACACATTTACTCATTC 1200
Oy 1201 GGTATCCCGCACCCGTGAAGGTAATATGTTGTTACTCAAGTTATATAGCAATAGAGC 1260
Db 1201 GGTATCCCGCACCCGTGAAGGTAATATGTTGTTACTCAAGTTATATAGCAATAGAGC 1260
Oy 1261 TTCTATCCAGAACATCACTCTCACCTGCGGACAAAGCTTGGGTTAATATTAAGGCTCT 1320
Db 1261 TTCTATCCAGAACATCACTCTCACCTGCGGACAAAGCTTGGGTTAATATTAAGGCTCT 1320
Oy 1321 GACACATCTGGAGAGAAATAGTTCCGGACACAGACAAATTTCCA 1365
Db 1321 GACACATCTGGAGAGAAATAGTTCCGGACACAGACAAATTTCCA 1365

```

RESULT 2

US-09-986-682b-5
Sequence 5, Application US/09986682B

Patent No. US20020115182A1
GENERAL INFORMATION:

APPLICANT: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU
KENKYUJO
TITLE OF INVENTION: POLYPEPTIDE HAVING - FRUCTOPURANOSIDASE
ACTIVITY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: BROMDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/986, 682B
FILING DATE: 06-NO. US20020115182A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 170,630/1996
FILING DATE: 10-JUN-1996

ATTORNEY/AGENT INFORMATION:
NAME: BROMDY, Roger L.

REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2408 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

ORGANISM: Bacillus sp.

INDIVIDUAL ISOLATE: V230 (FERM BP-5054)

FEATURE:
NAME/KEY: signal peptide
LOCATION: 1..360
IDENTIFICATION METHOD: E
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-986-682b-5

Query Match 100.0%; Score 1365; DB 10; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAACAGCGGGAGCTACAGAGAGACTATGTTGGCCATATTACAGCGCGTGACATG 60
DB 457 ATGAACAGCGGGAGCTACAGAGAGACTATGTTGGCCATATTACAGCGCGTGACATG 516
QY 61 CTAATAATTCAGAGCAACAACAGCTCAATTAAAGGCCCATTTCAATGACATCA 120
DB 517 CTAATAATTCAGAGCAACAACAGCTCAATTAAAGGCCCATTTCAATGACATCA 576
QY 121 GCAATCAAAAACATTTGATTCGGCAAAAAGGATATGTAAGTACAGCACTTAATAGATTYA 180
DB 577 GCAATCAAAAACATTTGATTCGGCAAAAAGGATATGTAAGTACAGCACTTAATAGATTYA 636
QY 181 GATGATGGGATAGCTGGCCACTGCAAAAACGCTGATGCTACTCGCGCAATTTATCATGA 240
DB 637 GATGATGGGATAGCTGGCCACTGCAAAAACGCTGATGCTACTCGCGCAATTTATCATGA 696
QY 241 TATACATCGTCTCGGCTTTAGCAGGTGACCCAAAACAGTATGATATCTCCATTCAT 300
DB 697 TATACATCGTCTCGGCTTTAGCAGGTGACCCAAAACAGTATGATATCTCCATTCAT 756
QY 301 TTATTCATCAAAAAGTCTGATATCATGATTCGACAGCTGGAATAATCTGGAAGATA 360
DB 757 TTATTCATCAAAAAGTCTGATATCATGATTCGACAGCTGGAATAATCTGGAAGATA 816
QY 361 TTTGAGATATGATTAATTTGTTCCAAATGATCCGTAATCTTAATATCAACACAGAG 420
DB 817 TTTGAGATATGATTAATTTGTTCCAAATGATCCGTAATCTTAATATCAACACAGAG 876
QY 421 TGGTCAGGTCTGCTACTTTAACCAAGATGCCAAGCGCTTATTCCTATACAGATATAC 480
DB 877 TGGTCAGGTCTGCTACTTTAACCAAGATGCCAAGCGCTTATTCCTATACAGATATAC 936
QY 481 TCAGATATCTGGAAGATGATGAACCGCTGCTGTAACCAATATTCACACTGCTCAA 540
DB 937 TCAGATATCTGGAAGATGATGAACCGCTGCTGTAACCAATATTCACACTGCTCAA 996
QY 541 GTAACTTATCCACGCCGATGACAGCTACACTTAAAGTCGATGATGATCAATAA 600
DB 997 GTAACTTATCCACGCCGATGACAGCTACACTTAAAGTCGATGATGATCAATAA 1056
QY 601 TCTGCTTTGATGGCGGAGAGGATACACTTATCAAAATATTCACCAATTTATCATGAA 660
DB 1057 TCTGCTTTGATGGCGGAGAGGATACACTTATCAAAATATTCACCAATTTATCATGAA 1116
QY 661 GGCAAGTGATTTGAGTGATACCATCTTAAAGAGCCCTCACTATGTGAAGATAG 720
DB 1117 GGCAAGTGATTTGAGTGATACCATCTTAAAGAGCCCTCACTATGTGAAGATAG 1176
QY 721 GGCCATAAATATCTTCTTTGAAGCGAATCTGAGAACAGAGATGTTATCAAGCGCAT 780
DB 1177 GGCCATAAATATCTTCTTTGAAGCGAATCTGAGAACAGAGATGTTATCAAGCGCAT 1236
QY 781 CAGCTTTTCATTAATTAAGCTTACATGCGGAGATGACGCTCTCTTCCAGATGAATAA 840
DB 1237 CAGCTTTTCATTAATTAAGCTTACATGCGGAGATGACGCTCTCTTCCAGATGAATAA 1296
QY 841 AATAAATGCTTCAAGTCTTAATAAACAATTTGCTTTTAAAGCAATGATGATAGGC 900
DB 1297 AATAAATGCTTCAAGTCTTAAAGTCTTAATAAACAATTTGCTTTTAAAGCAATGATGATAGGC 1356
QY 901 ATTGTGATTTGGCCGATGACTATACAGTGAAGAAGTGTATGAACAATTAAGTCGATCA 960
DB 1357 ATTGTGATTTGGCCGATGACTATACAGTGAAGAAGTGTATGAACAATTAAGTCGATCA 1416
QY 961 AACACAGTAGAGATGAGATGAGACGGCCCAATATTTAAATGAATAATTAATGAT 1020
DB 1417 AACACAGTAGAGATGAGATGAGACGGCCCAATATTTAAATGAATAATTAATGAT 1476
QY 1021 CTATTCACAGATATTCAGAGATATCCAAATATGACAGATGATGAATTAACGACAAAGAT 1080
DB 1477 CTATTCACAGATATTCAGAGATATCCAAATATGACAGATGATGAATTAACGACAAAGAT 1536
QY 1081 TATATGCTAGAGGCCCGGAGGAGACTCTTAATATGCGCACACACACCGATTAATGA 1140
DB 1537 TATATGCTAGAGGCCCGGAGGAGACTCTTAATATGCGCACACACCGATTAATGA 1596
QY 1141 GGACTTTATGAAATGATATCTTACCTGCTGATCCACACACACTTACTTCATTCG 1200
DB 1597 GGACTTTATGAAATGATATCTTACCTGCTGATCCACACACACTTACTTCATTCG 1656
QY 1201 GGTATCCCGCACCCCTGAAGATTAATATGCTACTCACACAGTTATATGACGAATAGAG 1260
DB 1657 GGTATCCCGCACCCCTGAAGATTAATATGCTACTCACACAGTTATATGACGAATAGAG 1716
QY 1261 TTTATCCAGAAATCACTCTCACCCTGCGGCAAGCTTTGGGTTAATATTAAGGCTCT 1320
DB 1717 TTTATCCAGAAATCACTCTCACCCTGCGGCAAGCTTTGGGTTAATATTAAGGCTCT 1776
QY 1321 GACACATCTGGAGGAGAAATATGTTCCGGACAGACAAATCCCA 1365
DB 1777 GACACATCTGGAGGAGAAATATGTTCCGGACAGACAAATCCCA 1821
```

RESULT 3

US-09-974-300-2644
Sequence 2644 Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 10085-500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2644
LENGTH: 1449
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-2644

Query Match 45.2%; Score 617.2; DB 10; Length 1449;
Best Local Similarity 67.3%; Pred. No. 4,2e-151;
Matches 896; Conservative 0; Mismatches 418; Indels 18; Gaps 1;

```

QY 12 GCATACACAGAGACATATGTTTGGCCATATTTACAGCGCTGACATGCTAAATATCC 71
DB 99 GGATTAACAAGAAACCTACAGATTTTCTCATATCAACAGACATGACATGCGAAATATCC 158
QY 72 AGGACACAAAACAGCTCTCAATTTAAAGGCTCAATTCATGATGATGACATCAAAA 131
DB 159 CGAGGAGCAAAAGACGCAACATTAATTAAGTTCTCAATTCATGCGAAACATCAAAA 218
QY 132 CATTCATTCGCGCAAAAGGATGATGAAGTCAAGCACTTAATATGATTAAGATGAGGA 191
DB 219 CATTCCTTCTGCAAAAGGATATACAAAATGAGAGCTGATCGATTTAGACGATGAGGA 278
QY 192 TAGCTGGCCACTGCAAAAACGCTGATGATGCTGCGCAATTAATCATGATATCATCATCGT 251
DB 278 TAGCTGGCCACTGCAAAAACGCTGATGATGCTGCGCAATTAATCATGATATCATCATCGT 340
```

Db 279 CAGTGGCCGCTGCAAAATGCCAGGAGCGGTCTGTACATACCAAGCTTACTTGT 338
OY 252 CTCGGCTTTAGAGGAGGACCAAAAAACGATGATGATCTACCTCATTTTTCATCA 311
Db 339 TTTTGGCTGGGGGATCGCAAAAGCTGATGATGATGATGATGATGATGATGATGAT 398
OY 312 AAAAGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
Db 399 AAAAGAGGCGAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 458
OY 372 GATTAATTTTGTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
Db 459 CGACAAATTTTGTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 518
OY 432 TGTCTATTAAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491
Db 519 TGGCAGCGTACAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 578
OY 492 TGAAGATGGTGAACCGGCTGTGTAAACCAATTCATTTCACTGCTCAAGTAACTTATC 551
Db 579 ATA-----CGGCAAGCAGACGCTGACACAGCTCAGTCAATTTCTC 620
OY 552 CCAGCCGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
Db 621 TGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
OY 612 TGGCGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Db 681 CGGCGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 740
OY 672 TTTAGTGTATACCAATCTTAAAGACCTTAAAGATGATGATGATGATGATGATGATGAT 731
Db 741 CTCGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 800
OY 732 TCTTGTCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
Db 801 TCTTGTCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 860
OY 792 TATAAAGCTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
Db 861 CAACAGAGCTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 920
OY 852 TCAAGTCTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
Db 921 GCAAGTCTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980
OY 912 GCGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
Db 981 AATTAAGCTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
OY 972 AGATGAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
Db 1041 AATTAAGCTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100
OY 1032 TTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
Db 1101 TTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160
OY 1092 GCGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
Db 1161 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220
OY 1152 GAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1211
Db 1221 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1280
OY 1212 CCTTGAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1271
Db 1281 GAAAAAGGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1340
OY 1272 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1331
Db 1341 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1400

OY 1332 AGCAAAATAG 1343
Db 1401 TGTCAAAACAG 1412

RESULT 4
US-09-993-292A-19
Sequence 19, Application US/09993292A
Patent No. US20020146430A1
GENERAL INFORMATION:
APPLICANT: James E. Galen
APPLICANT: University of Maryland
TITLE OF INVENTION: USE OF CLY A HEMOLYSIN FOR EXCRETION OF
FILE REFERENCE: JOMED.007A
CURRENT FILING DATE: 2001-11-23
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 2253
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (0)...(2253)
US-09-993-292A-19

Query Match
Best Local Similarity 67.28; Score 578.6; DB 10; Length 2253;
Matches 903; Conservative 0; Mismatches 389; Indels 51; Gaps 4;

OY 1 ATGAACAGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 922 ACGAACAACCAACCTATATGAGAAACATACGCGATTTCCATATTAACAGCGCATGATGAT 981
OY 61 CTAAATTTCCAGGACACAAACAGCTCTCATTTAAAGTCCCATTTCAATTCATCA 120
Db 982 CTGCAATTCCTGAAACAGCAAGCAAAATATGCAAGTTCGAAATTCGATTCGTC 1041
OY 121 GCAATCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 1042 ACAATTTAAATATCTCTTCTGCAAAAG-----CCTG 1074
OY 181 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 1075 GACGTTTGGGACAGCTGGCATTAACAAACCTGACGCGACCTGCAAACTATACAGGC 1134
OY 241 TATCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 1135 TACCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
OY 301 TTTATCTATCAAAAGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 1195 AGTCTCTATCAAAAGTGGGCAAACTTCTATGACAGCTGCAAAAGCTGGCGCTC 1254
OY 361 TTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 1255 TTTAAGAGAGGCAAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
OY 421 TGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 1315 TGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
OY 481 TCAAGTAAATCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 1375 TCGGTAACATTA-----CGGCAAAACAAACCTGCAACATCTGACAA 1416
OY 541 GTAACCTATCCACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

QY 1321 GACACATCTGAGAGAAAAATAG 1343
 Db 2871 AAAACATCTGTGTCAAAAGACAG 2893

RESULT 7

US-10-055-794-3
 ; Sequence 3, Application US/10055794
 ; Publication No. US20030059900A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Farmer, Alan Andrew
 ; TITLE OF INVENTION: SEQUENCE SPECIFIC RECOMBINASE-BASED
 ; TITLE OF INVENTION: METHODS FOR PRODUCING INTRON CONTAINING VECTORS AND
 ; FILE OF INVENTION: COMPOSITIONS FOR USE IN PRACTICING THE SAME
 ; CURRENT APPLICATION NUMBER: US/10/055,794
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/263,358
 ; PRIOR FILING DATE: 2001-01-18
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 6525
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Vector
 US-10-055-794-3

Query Match 42.3%; Score 577; DB 9; Length 6525;
 Best Local Similarity 67.2%; Pred. No. 3,1e-140;
 Matches 902; Conservative 0; Mismatches 390; Indels 51; Gaps 4;

QY 1 ATGACAGGGGGGATACAGAAGACATATGTTTCCCATATTAACAGCGCTGACATG 60
 Db 3189 ACGAACCAAAACCATATTAAGAAACATAGCGCATTTCCCATATTAACAGCGCATATG 3248
 QY 61 CTAAATATTCAGACACAAACAGTCCCTCAATTTAAAGTGCCTCAATTCAGTCA 120
 Db 3249 CTGCAATCCCTGAGACAGCAAAATGAAATATCAAGTTCGAGTTCATTCGTC 3308
 QY 121 GCAATCAAAACATTTGATGCGCAAAAGGATATGATAGTCAGGCACTTAATAGATT 180
 Db 3309 ACAATTTAAATATCTTCTGCAAAAG-----CCG 3341
 QY 181 GATGATGGGATAGTGGCCACTGCAAAACGCTGATGACTGCGCAATATATGATGA 240
 Db 3342 GACGTTTGGGACAGCTGGCCATTACAAAACGCTGACGGCAGTGTGCAAACTATGACGCG 3401
 QY 241 TATCATCGTCTCGCTTTTACAGGTGACCCAAAACAGATGATATCTCCATCAT 300
 Db 3402 TACCATCGTCTTGGCATTTAGCCGAGATCTTAAATATGCGGATGACATGATTTAC 3461
 QY 301 TTATTTATCAAAAAGTGGTGATCATGATGATGCAAGTGAATAATGCTGAAGAGTA 360
 Db 3462 ATGTTCTATCAAAAAGTGGCGAACTTCTATTTGACAGCTGAAAAACGCTGGCGGCTC 3521
 QY 361 TTTGAAGATATGATTAATTTGTTCGAATGATCCGTATCTTAATATCAACACAGAG 420
 Db 3522 TTTAAAGACAGGAGCAAAATTCGATGCAATATGATTTATCTTAAAGACCAACAGAA 3581
 QY 421 TGTGAGGTCTGCTACTTTTAAACCAAGATGCGCAAGTCCGTTTATTCATACAGATT 480
 Db 3582 TGTGAGGTCTGCTACTTTTAAACCAAGATGCGCAAGTCCGTTTATTCATACAGATT 3641
 QY 481 TCAGGTAATCTGGAAGATGTGGAACGCGTGTGTAACCAATCATTTCAACTGCTCA 540
 Db 3642 TCCGTTAAACATTA-----CGCAACAAACACACTGACACTGACACAA 3683
 QY 541 GTAACCTATCCAGCCGATGACACTACATTAAGTCGATGAGTATCTATATATAA 600
 Db 3684 GTTAAGTATCAGCATGACAGAGCTCT--TTGAACATCAACGAGTGTAGAGATTATATAA 3740

QY 601 TCTGCTTTGATGGCGAGACGCTACAGTTTATCAAAAATATTCAGCAATTTATCGAGAA 660
 Db 3741 TCAATCTTTGA---CGTGACGCGAAAAAGATATCAAAATGTACAGAGTTCATCATGA 3797
 QY 661 GGCAGTGGATTTTCAGGTATTAACCATCTTTAAGAGACCTCACTATGTGGAAGATAAG 720
 Db 3798 GGCACCTACAGCTCAGGCGACCAACCATACGTGAGATCCTCACTAGTACAGTAA 3857
 QY 721 GGCATTAATATCTGTCTTTAAGCGAATACTGGAACACAGATGTTATACAGCGAT 780
 Db 3858 GGCACAAATATCTGTATTTAGCAAAACCTGGAACGAGATGGCTACCAAGCGAA 3917
 QY 781 CAGTCTTCAATATAAACCTTACTATGCGGAGAGGAGCTCTCTTCAGAAATGAAA 840
 Db 3918 GATCTTTTATTAACCAACATACATCTGCAAAAGACATCTTCTCCGTCAAGAACT 3977
 QY 841 AATAAAGCTCTCAAAAGTCTAAAAACAAATTTGCTTTTAAAGCAATGGTCATTAGCC 900
 Db 3978 CAAAAACCTCTGCAAAAGCGATTAATAAAGCGCGTGAAGTTGCAAAAGCGCGCTCGGT 4037
 QY 901 ATTGTTGATTTGGCGGATGACTATACAGTGAAGAGTGTATGAACCATTAAGCTCA 960
 Db 4038 ATGATTTAGCTAAAGCATGATTAACACACTGAAAAAGTATGAACCGCTGATTCATCT 4097
 QY 961 AACACAGTAGCAGATGAAGTGAAGCGCCCAATATATTTAAATGAATATATGAT 1020
 Db 4098 AACACAGTACAGATGAATTTGAACCGCGCAACGCTTTTAAATGAACGCAATATGTC 4157
 QY 1021 CTATTCAGCGATTCAGAGATCCAAATATGACAGATGATGAATTAACGACAAAGATGT 1080
 Db 4158 GTTTCAGTACGATCCCGGATCAAAATGACGATTTAGCGGATTAAGCTACGCTAAGCTAT 4217
 QY 1081 TATATGCTAGGCGCGGAGGCGACATCTTAATGCGCCACACACCCGTAATGAAT 1140
 Db 4218 TACATGCTGTGATATGTTCTAATCTTAACTGACCCATTAACCCGCTGAACAAACT 4277
 QY 1141 GCACTGTATTAACATGATCTTGAACCTGCTGATCTACACACACTACTCTCATGTC 1200
 Db 4278 GGCCTGTGTAAATGATTTGATCTTACAGATGATACCTTTTACTTACTACACTTC 4337
 QY 1201 GGTATCCGACCCCTGAAGTATATGATGCTACTACAAAGTTATATGACGATGAGGC 1260
 Db 4338 GCTGTACCTCAAGCAAAAGCAATGATGATGATTAACAGCTATATGCAACAAAGAGA 4397
 QY 1261 TTCTATCCAGACATCACTCTCACTGCGGCAACGTTGGGTTAATTAAGGCTCT 1320
 Db 4398 TTCTACGACAGCAAAACATCAAGTGTGGCCTAGCTTCTGCTGAACATCAAAAGCAAG 4457
 QY 1321 GACACATCTGGAGAAAAATAG 1343
 Db 4458 AAAACATCTGTGTCAAAAGACAG 4480

RESULT 8

US-09-849-866-1/C
 ; Sequence 1, Application US/09849866
 ; Patent No. US2002006833A1
 ; GENERAL INFORMATION:

APPLICANT: Ilya Chumakov
 Hiroaki Tanaka
 TITLE OF INVENTION: High Throughput DNA Sequencing Vector
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:

ADDRESS: Martens, Olson & Bear, LLP
 STREET: 550 West C Street, Suite 1200
 CITY: San Diego
 STATE: California
 COUNTRY: USA

ZIP: 92101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: win95
SOFTWARE: word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/849,866
FILING DATE: 04-MAY-2001
ATTORNEY/AGENT INFORMATION:
NAME: Daniel Hart
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.15CDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10317 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: CIRCULAR
MOLECULE TYPE: synthetic DNA
ORIGINAL SOURCE:
ORGANISM: cloning vector pGendel
FEATURE:
NAME/KEY: pGendel
LOCATION: 1..10317
FEATURE:
NAME/KEY: Homology with X06404 comp1 (411..1668)
LOCATION: 9..1266
IDENTIFICATION METHOD: blastn against X06404
FEATURE:
NAME/KEY: Kanamycin resistance gene CDS
LOCATION: 142..957
IDENTIFICATION METHOD: By homology to X06404
FEATURE:
NAME/KEY: Tn1000's right end
LOCATION: complement 1332..1371
IDENTIFICATION METHOD: blastn against X60200)
FEATURE:
NAME/KEY: Homology with U46017 (1-472)
LOCATION: 1423..1894
IDENTIFICATION METHOD: blastn against U46017
FEATURE:
NAME/KEY: single stranded DNA replication origin
LOCATION: 1423..1894
IDENTIFICATION METHOD: By homology to U46017
OTHER INFORMATION: mutation T -> C 1658
FEATURE:
NAME/KEY: Homology with U51113 (2382..6997)
LOCATION: 1896..6544
IDENTIFICATION METHOD: blastn against U51113
FEATURE:
NAME/KEY: OriS
LOCATION: 1972..2188
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: repELR
LOCATION: 2897..2918
OTHER INFORMATION: Described in seqid 16
FEATURE:
NAME/KEY: RepE
LOCATION: 2903..3034
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: T3
LOCATION: 3043..3059
OTHER INFORMATION: Described in seqid 17
FEATURE:
NAME/KEY: LRT3RA
LOCATION: complement 3045..3069
OTHER INFORMATION: Described in seqid 15
FEATURE:
NAME/KEY: IncC
LOCATION: 3070..3320
IDENTIFICATION METHOD: By homology to U51113

OTHER INFORMATION: insertion 33 bases 3038..3071
FEATURE:
NAME/KEY: ParA
LOCATION: 3655..4821
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: mutation G -> A 3878
FEATURE:
NAME/KEY: ParB
LOCATION: 4821..5792
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: ParC
LOCATION: 5865..6382
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: Homology with J01688 (complement 175..819)
LOCATION: 6574..7218
IDENTIFICATION METHOD: blastn against J01688
OTHER INFORMATION: mutation A -> G 7096
FEATURE:
NAME/KEY: CDS streptomycin sensitivity gene
LOCATION: complement 6716..7090
IDENTIFICATION METHOD: By homology to J01688
OTHER INFORMATION: mutation A -> G 6728
mutation G -> C 6821
mutation C -> T 6866
mutation T -> C 7013
mutation T -> A 7058
FEATURE:
NAME/KEY: rpsLr
LOCATION: 7155..7174
OTHER INFORMATION: Described in seqid 12
FEATURE:
NAME/KEY: SP6
LOCATION: 7230..7248
OTHER INFORMATION: Described in seqid 13
FEATURE:
NAME/KEY: Tn1000's left end
LOCATION: 7252..7291
IDENTIFICATION METHOD: blast (X60200)
FEATURE:
NAME/KEY: Homology with X02730 (complement 37..1959)
LOCATION: 7305..9227
IDENTIFICATION METHOD: blastn against X02730
FEATURE:
NAME/KEY: CDS levansucrase gene
LOCATION: complement 7379..8800
IDENTIFICATION METHOD: By homology to X02730
OTHER INFORMATION: mutation T -> C 7466
mutation A -> G 7739
mutation T -> C (Asn -> Asp) 8347
mutation T -> C 8600
mutation G -> A (Ala -> Val) 8772
FEATURE:
NAME/KEY: SLR3
LOCATION: 8711..8731
OTHER INFORMATION: Described in seqid 14
FEATURE:
NAME/KEY: Homology with J01636 (complement 1158..1465)
LOCATION: 9298..9623
IDENTIFICATION METHOD: blastn against J01636
FEATURE:
NAME/KEY: CDS alpha peptide beta-galactosidase
LOCATION: complement 9276..9497
IDENTIFICATION METHOD: By homology to J01636
FEATURE:
NAME/KEY: primer HE1
LOCATION: complement 9465..9479
FEATURE:
NAME/KEY: primer HE2
LOCATION: 9461..9475
FEATURE:
NAME/KEY: primer LacLR52AVT


```

NAME/KEY: unsure
LOCATION: 362
NAME/KEY: unsure
LOCATION: 363
NAME/KEY: unsure
LOCATION: 364
NAME/KEY: unsure
LOCATION: 365
NAME/KEY: unsure
LOCATION: 366
NAME/KEY: unsure
LOCATION: 367
NAME/KEY: unsure
LOCATION: 368
US-09-864-761-21746

Query Match
Best Local Similarly 64.0%; Score 135.6; DB 10; Length 368;
Matches 235; Conservative 0; Mismatches 105; Indels 27; Gaps 1;

16 TACAAAGAACACATATGTTTGGCCATATTACACGGCGCTGACATGCTAAATAATTCACAGA 75
|||||
12 TACAAAGAAACGTAACGGCGCTCTCTCATATTACAGCGCATATATGCTGCGATCCCTAA 71
|||||

76 CAACAAACAGCTCCTCAATTAAAGTGCCCTCAATTCAATGCATCAGCAATCAAAACATT 135
|||||
72 CACAGACAAACGAAATAATACCAAGTGCCCTCAATTGATCAATCAACGATTTAAATAATT 131
|||||

136 GATTGGCAAAAGGATGATTAAGTCAGCAACTTAATTAAGTATGATGATGATGATG 195
|||||
132 GAGCTGTGCAAAAGGACTT-----GATGTGTGGGACAGC 164
|||||

196 TGGCCACTGCAAAACGCTGATGTACTGCGGCAAAATTATCATGATATCATCTCTCC 255
|||||
165 TGGCCGCTGCAAAACGCTGACGAGACAGTAGTGAATATACAACGGCTATCACGTTGCTTT 224
|||||

256 GCTTTAGCAGGTGACCCAAATAAACAGTGTGATGATCTCACTTATTTATTTCAAAA 315
|||||
225 GCTCTTGGCGGGAAGCCCGGAAGACGCTGATGACACATCAATCAATGTTTATCAAAAG 284
|||||

316 GTGCGTATCATGATGATTCAGACCTGGAATAATGCTGGAAGATATTTGAAGATATGAT 375
|||||
285 GTGCGGCAACAATCAATGATGACACCTGGAATAAACGCGGCGGTCTTTNTTAAGACAGCGAT 344
|||||

376 AAATTG 382
|||||
345 AAGTTCG 351
|||||

RESULT 10
US-09-864-761-5018
; Sequence 5018, Application US/09864761
; Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27

```



```
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 5018
;; LENGTH: 376
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
```

```
;; OTHER INFORMATION: MAP TO AL035452.6
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
;; US-09-864-761-5018
```

```
Query Match 5.4%; Score 73.2; DB 10; Length 376;
Best Local Similarity 62.2%; Pred. No. 3.4e-09;
Matches 148; Conservative 0; Mismatches 63; Indels 27; Gaps 1;
```

```
QY 16 TACAAGGAAGACTATGTTTGGCCATTTTACAGCGCGTACATGCTAAATTCACAGA 75
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163 TACAAGGAAGACTATGTTTGGCCATTTTACAGCGCGTACATGCTAAATTCACAGA 222
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 CAACAAGAGCTCTCTCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAACATT 135
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 CAGCAGCAAAAGCAAAATACCAAGTGCCTCAATTCATGATCAGCAATCAAAATATT 282
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 GATTTCGCAAAAGGATGATAGTACAGCAACTTAATAGATTGATGATGATAGC 195
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 GAGTCTGCAAAAGGACTT-----GATGTGGGAGCAGC 315
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 TGGCCACTGCAAAAGCTGATGCTACGCGCAATTTATCATGATACATGCTCT 253
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 TGGCGCGCTGCAAAAGCTGACGAGCAACAGTACGATGATACACGCGTATCAGCTTGTGT 373
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-09-778-844-136
; Sequence 136, Application US/09778844
; Patent No. US20020150971A1
; GENERAL INFORMATION:
; APPLICANT: JOHANSEN, JEANETTE ELISABETH
```

```
;; APPLICANT: SCHALLING, MARTIN
;; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
;; FILE OF INVENTION: INTAKE AND/OR BODY WEIGHT
;; FILE REFERENCE: 030307/0195
;; CURRENT APPLICATION NUMBER: US/09/778,844
;; NUMBER OF SEQ ID NOS: 206
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 136
;; LENGTH: 1146
;; TYPE: DNA
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: CPLA2-beta, EMBL
;; US-09-778-844-136
```

```
Query Match 5.3%; Score 72.4; DB 10; Length 1146;
Best Local Similarity 66.9%; Pred. No. 1e-08;
Matches 103; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
```

```
QY 16 TACAAGGAAGACTATGTTTGGCCATTTTACAGCGCGTACATGCTAAATTCACAGA 75
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 954 TACAAGGAAGACTATGTTTGGCCATTTTACAGCGCGTACATGCTAAATTCACAGA 1013
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 CAACAAGAGCTCTCTCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAACATT 135
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1014 CAGCAGCAAAAGCAAAATACCAAGTGCCTCAATTCATGATCAGCAATCAAAATATT 1073
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 GATTTCGCAAAAGGATGATAGTACAGCAACT 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1074 GAGTCTGCAAAAGGACTGATGCTGTGGACAGCT 1107
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 12
US-09-995-587A-10
; Sequence 10, Application US/0995587A
; Patent No. US20020127681A1
; GENERAL INFORMATION:
; APPLICANT: VAN HUIJUM, SACHA ADRIANUS FOKKE TACO
; APPLICANT: VAN GEEL-SCHOUTEN, GERRITDIANA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; TITLE OF INVENTION: NOVEL FRUCTOSYLTRANSFERASES
; FILE REFERENCE: BO43667-CIP
; CURRENT APPLICATION NUMBER: US/09/995,587A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 09/604,958
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201872.9
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4634
; TYPE: DNA
; ORGANISM: Lactobacillus reuteri
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1220)..(3598)
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1205)..(1210)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2702)..(2707)
; OTHER INFORMATION: a, c, t, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (3686)..(3698)
; OTHER INFORMATION: a, c, t, g, other or unknown
; US-09-995-587A-10
```

OY		254	CCGCTTACAGCGTGACCACCAAAAAACGTATGACTCCACTTCATTATTTCTATAA	313
Db		854	TCCGAATGATGGGAATTCCAAACCAA --- ATGATAATCATATCTATCTTTATATA	910
OY		314	AAGTCGSGTAATACATGCATGTAGCAGCGTGGAAAAAGTGTGAGAGATTTGGAAATATGG	373
Db		911	AGATVGGTGAAATGAATTAAATGATCATTTGGAAGAATGATAGTCCAAATTTTGGCTATATTT	970
OY		374	ATAAATTTGTTCCAAAATGATCCGTAATCTTAAATATCAACAACAGSAGTGGCTAGTTCG	433
Db		971	CTACCCGGGTT ----- TCACAAGAATGGTCAAGATCAC	1003
OY		434	CTACTTTAACCAAGATGCCAAGTCGCTTATTTCTATACAGATTACTCAGTTAATCTG	493
Db		1004	CTGTTTGGAACAGATGATACATCTATCCAAATTAATTTATFACA ----- AGGDTAG	1051
OY		494	AAGATGSGTAGAOCGGGTGCTGGTATACCAATCAATTCOAATCGTCCAAGTAACATTATCCC	553
Db		1052	ACACGCTGATATCAATATCAACATCATCAAAAATATGCTACGGCTACTCTTATTTATAC	1111
OY		554	AGCCGATGACAGCTACACTTAAAGTCGATGGAGATGTGATCATCAATCTGTCTTGATG	613
Db		1112	ATAATATGGAATATGATCTACTGCTCAGGTACGAATATGACTATA ----- TGTATATT	1165
OY		614	GCGGAGACGGGTACAGTTATTCAAAATATTCAGCAATTTATCATGANGAAGCAAGTGATTT	673
Db		1166	AAGSTGATGGCTATTACTACCAAACTTATGATCATGGAAGGCTACTAACAAAGG --- TG	1222
OY		674	CAGSTGATTAACATNACTTTAAGAGACCCCTCAATATGTTGAAGATMAAGGCCATTAATATC	733
Db		1223	CCGATTAATATTCGATCGCTGATGCTCATATTAATTGAAGATGGTAATGGATGGATCGGATC	1282
OY		734	TTGCTTTGAAGCAATACTGSAACAACAGATGGTTATCAAGCGATCACTGCTTTCATA	793
Db		1283	TTGTTTTTGAAGCAAGTACTGG --- TTGSAATAATTTCAAGGCGAGGACCAATTTATA	1339
OY		794	ATAAAGCTTACTATGGCGGAGTGCAGCTCTTCTTCCGAATGAAAAATAACTGCTTC	853
Db		1340	ACTGTTAAATTAATGAGCGGAGATGACCATTTAATATCAAGAGCTTATTTAGAAATCTT	1399
OY		854	AAAGTCTTAAAAACAANNT --- GCTTCCTTAAAGCATGSGCATTTGSCATTTGGTAAT	910
Db		1400	CCAATGATGATTTAAGATGCGGCAACTTGGGCTAATGCAAGCTATCGTATCCTCAAC	1459
OY		911	TGCGCGATGACTA ----- TACAGTGAANAATGTTATGAAAACCTATAGTGCATCAA	961
Db		1460	TAAATTAAGAGCAAAACATCCTTAAGGTGGCAGAGTTATATCTCACCTTAATTTCTGAC	1519
OY		962	ACACAGTAGCAGATGAAGTCGAACGCCCAATATATTTAAATGAATGAATAAATGCTATC	1021
Db		1520	CAATGTAAGAGGATGAATGAGGACCAATATGAGTTAATTAAGTAATTAATTAATTA	1579
OY		1022	TATTCACGGGATTCGAAGAGATCCAAAATGACAGATGTSANATTAACGACAAAGATGTT	1081
Db		1580	TATTTGGCGGTACCCGTTTAAATCGAAGAATGATATGATGCTTGGATGAATGCTAATT	1639
OY		1082	AT ----- ATGCTAGGGCCGAGGCGACCTCCTTAATAGGCCAC	1120
Db		1640	ATGCGCTTGGTGAATAGTTGECANTGGTGGATATGTTGCTGATAGTCTTACTGEGATCTT	1699
OY		1121	ACAACCCGATTAATGAAGAACTGAGCTTGTATTTGAACATGAATCTTTCACCTCGTCA	1180
Db		1700	ATAAGCATTAATATGATTTCTGAGATATCTTGAATCTGTTCTGTTCCGCAAACTGGCGGA	1759
OY		1181	CACACACTTACTCATTCGCGGTATCCGCAACCTG --- AAGSTAATATGTGTA	1237
Db		1760	CAGCAACTTATTCATATTAATGCTGTCCCGCTTGGCGGAAGAATGACCAAGATTTAGTTA	1819
OY		1238	CMACTTATATGACGAATAGA	1257
Db		1820	CTTCATATATGACTAATAGA	1839

```

RESULT 14
US-09-995-587A-3
; Sequence 3, Application US/09995587A
; Patent No. US2002127681A1
; GENERAL INFORMATION:
; APPLICANT: VAN HIJUM, SACHA ADRIANUS FOKKE TACO
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDIJNA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; TITLE OF INVENTION: NOVEL FRUCTOSYLTRANSFERASES
; FILE REFERENCE: BO43667-CIP
; CURRENT APPLICATION NUMBER: US/09/995,587A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 09/604,958
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201872.9
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Lactobacillus reuteri
US-09-995-587A-3

Query Match      3.8%; Score 52; DB 10; Length 2394;
Best Local Similarity 48.3%; Pred. No. 0.0032;
Matches 589; Conservative 0; Mismatches 535; Indels 96; Gaps 12;

OY      80  AAACAGTCCTCAATTTAAAGTGCCTCAATTCATCATCAGACATCAAAACATTTGAT 139
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      701  AACAGATGTCGTGACACAGTTCATCTTTAAAGCAAGTAATCAAAATATGCGCG 760

OY      140  CGGCAAAAGGTATATAGTACGACCACTTAATGA---TTAATGTATGGGTAGCT 196
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      761  CGGTCACACTTAAAGTACCAAACTAATCTATTGAACCTTTAATGATGCGATTTAT 820

OY      197  GGCACACGCAAAAGCCT---GATGCTACTGCGCAAAATTTATGATATGATATGATGCT 253
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      821  GGCAGCTTCAAGATTTTGGACAGACCAAGTTGCTAATTTGATATGCTATGCACTTGCA 880

OY      254  CCGCTTTAGCAGGTGACCCCAAAAAACAGTATGATCTCCACTTCATTTATTCATCAA 313
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      881  TCGCATATATGGGAATTCCAAAACAA---ATGATATATCATATCTATCTTATATATA 937

OY      314  AAGTGTGATATCATCATGATGACAGTGAATAATGCTGGAAGATATTTGAGATGCG 373
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      938  AGTATGTATATGAATTAAGTCAATTTGGAAGATGTAGTCCAAATTTTGGCTATTAAT 997

OY      374  ATAAATTTGTTCCAAATGATCCGTATCTTAATATCAAAACAGAGAGTGTGAGTTCG 433
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      998  CTACCGCGGT-----TCACAAAGATGTAGATCAG 1030

OY      434  CTACTTTAACCAAGATGGCAAGTCGCTTATTTCTATACAGATTACTAGTAACTATCC 493
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1031  CTGTTTGAACAGTATCACTATCTATCAATATTTTATACA-----AGGTAAG 1078

OY      494  AAGATGTGGAACCGGTGCTGGTACCAATTCATTCAACGCTCAAGTAACTATCC 553
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1079  ACACGCTCTGATTAACATACCAATTCATCAAAAATTTGAGCGCTACTCTTTATTAACG 1138

OY      554  AGCCGATGACAGTCACTTAAGAGTGAATGATGATCATCATTAATCTGCTTTGATG 613
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1139  ATAAATGAAGAAATGTATACATCGCTCAGTACGAAGAAAGACATAT-----TTGATATTTG 1192

OY      614  GCGGAGACGTACATTTATCAAAATATTCAGCAATTTATGATGAAGCAAGTGAATTT 673
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1193  AAGGTGATGCTATTTACTACCAAACTTATGATCAATGGAAGCTACTAACAAG---TG 1249

OY      674  CAGGTATACCATTTCTTAAGAGACCCCTACTATGTTGAAGATTAAGGCCCTAATATAC 733
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1250  CCGATATATTTGCAATGCGGTGATGCTCATGTAATTTGAAGATGTATAGTGTAGTAC 1309

```

```

OY      734  TTGTCTTGAACGAATATCGAACACAGATGTATCAAGCGCATCAGTCTTTCATA 793
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1310  TTGTTTTCACCAAGTACTG---TTTGAAATAATTCACAGGACCAAGCAAAATTTATA 1366

OY      794  ATAAAGCTTACTATGCGGAGTACGCTTCTTCCAGATGCAAAAAATAAATGCTTC 853
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1367  ACTGTTAAATTAATTTATGCGGAGATGACCGCATTTAATATCAAGAGCTTATTAATTAATTC 1426

OY      854  AAGTCCATAAAACAAAT---GCTTCTTTAGCAATGCTCATTTAGCATTTGAT 910
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1427  CCAATGATGATATTTAAGAGTGGCACTTGGGCTAATGACACTATCGGTATCCCAAC 1486

OY      911  TGGCCGATGACTA-----TACAGTGAAGTGTATGAACCATTAAGTCGATCA 961
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1487  TAAATAGACCAAAAGATCTTAAGGTGCGAGATTAATACACCATTAATTTCTGCAC 1546

OY      962  ACACAGTACAGATGAAGTCAAGCGCCCAATATTTAAATGAATATAAATGATGATC 1021
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1547  CAATGTTAAGCGATGAATTAATGACGACCAAAATGATGATTAATTAATTAATTAAT 1606

OY      1022  TATTACGATTCAGAGATTCACAAATGACGAGTATGATTAACGACCAAGATGTT 1081
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1607  TATTTGCCGTACCGCTTTAAATCGAGGAAGTAATGATGATGCTTGGATGAATGCTAAT 1666

OY      1082  AT-----ATGCTAGGCGCCGAGGCGCATCTTAATTTGCCCCAC 1120
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1667  ATGCGGTGATATATGTTGCAATGTCGGAATATGTTGCTATATGCTATCACTGATCTT 1726

OY      1121  ACAACCGATTAATGAACACTGACTTGTATTAAGATGAATCTGACCTGCTGATCA 1180
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1727  ATAAACCATTAATGATTTGAGTAGTACTGACAGTCTGCTGCTGCTGCTGCTGCTGCTG 1786

OY      1181  CACACTTACTCTATTTCCGATATCCGACCCCTG---AAGTAAATATGCTGACTCA 1237
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1787  CAGCAACTTATCATATTAATGCTGTCCCGTTGCCGAAAGATGACCAAGTATTAATTA 1846

OY      1238  CAAGTATATGACGAATAGA 1257
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1847  CTTCATATATGACTAATAGA 1866

```

```

RESULT 15
US-09-995-587A-4
; Sequence 4, Application US/09995587A
; Patent No. US2002127681A1
; GENERAL INFORMATION:
; APPLICANT: VAN HIJUM, SACHA ADRIANUS FOKKE TACO
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDIJNA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; TITLE OF INVENTION: NOVEL FRUCTOSYLTRANSFERASES
; FILE REFERENCE: BO43667-CIP
; CURRENT APPLICATION NUMBER: US/09/995,587A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 09/604,958
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201872.9
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Lactobacillus reuteri
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(51)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (68)..(2434)
US-09-995-587A-4

```

Query Match 3.8%; Score 52; DB 10; Length 2592;
 Best Local Similarity 48.3%; Pred. No. 0.0034;
 Matches 589; Conservative 0; Mismatches 535; Indels 96; Gaps 12;

```

OY 80 AAAACAGTCCTCAATTTAAAGTCCTCAATTCATATGATCAGCAATCAAAAACATTGATT 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 741 AACAGATGTGTGGTACACAGTTCATTTAAAGCAGTGAAATCAAAAATATGCGCG 800
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 140 CGGCAAAAGGTATGATGATGAGCAGCACTTAATAGA---TTAGATGATGGATAGCT 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 801 CCGCTACACTAAAGATGACAACTAATATCTATGTAACCTTTAGATGATGGATTCAAT 860
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 197 GGCACCTGCAAAACGCT---GATGGTACTGCGGCAATTATGATGATATCAATCGTCT 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 861 GGCAGATTCAGATGTGTGGACAGCAAGTTGTAATGGAATGGCTATCAACTGTCA 920
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 254 CCGCTTTACAGGTGACCCCAAAAACAGATGATATCTCACTTCATTATTCATCAAA 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 921 TCGCAATGATGGAAATCCCAACAAA---ATGATATATATCTATCTATATATAA 977
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 314 AAGTCGTGATCATGATGATGACAGTGGAAAAATGCTGGAAGATTTGAAGATATGG 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 978 AGTATGATATATGATATGATTAATGATCATGAGAAATGTAGTCCAAATTTTGGCTATATT 1037
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 374 ATAAATTGTTCCAAATGATCCGTAATCTTAATATCAAAACAGAGAGTGTGAGTTCTG 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1038 CTACCGCGGTT-----TCACAGAAATGTCAGGATCAG 1070
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 434 CTACTTTAACCAAGATGGCCAGTCCGTTTATTCATACAGATTAATCTCAGTAATCCTG 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1071 CTGTTTGAACGTGTAATCTATTCCAATTAATTTATACA-----AGGGTAG 1118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 494 AAGATGTTGGAACCGTGTGCTGTAACCAAAATCATTTCACTCAAGTAAACTATATCC 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1119 ACACGCTGATTAACATATACCATCATCAAAAATTTGCTAGCGTACTCTTTATTTAATCG 1178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 554 ACCCGATGCACTACACTTAAAGTGTGATGAGTATCTGATCAATAATCTGCTTTGATG 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1179 ATAAATATGGAATGATCATCTCGCTCAGGTACGAATGACTATA-----TTGATTTTG 1232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 614 GCGGAGAGGTACAGTTTATCAAAATATTCAGCAATTTATGATGAGGCAAGTGTATT 673
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1233 AAGGTGATGTCTATTACTACCAACTTATGATCAATGGAAGTACTACTAAAGG---TG 1289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 674 CAGGTGATTAACCTACTTTAAGAGACCTCAGTATGTTGAAGATTAAGGCGCATTAATATC 733
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1290 CCGATAATATTCATATGCGTGTATGCTCATGTAATGGAATGATATGATGATCGGTACC 1349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 734 TTGCTTTGAACGATACTGGAACAACAGATGTTATCAGGCGATCAGTCTTTCATA 793
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1350 TTGTTTTTGAACAGTACTGTG---TTTGAAAAATTTATCAAGGCGAGCAAAATTTATA 1406
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 794 ATAAAGCTTACTATGCGGAGTGAAGTGTCTTCCAGATGAATAAATAAATGCTGTC 853
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1407 ACTGTTAATTTATGCGGAGATGAGCACTTAATATTCAGAGCTTATTTAGAAATCTTT 1466
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 854 AAAGTCTTAAAAACAATTT---GCTTTTAGCGATGTGATGATGATGATGATGATGAT 910
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1467 CCAATGATGATATTTAAGATGCGGCACTGGGCTAATGCACTATCGGTATCCTCAAC 1526
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 911 TGCGCGATGACTA-----TACAGTAAAAAGTGTATGAAACCATTAGTGCATCA 961
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1527 TAAATAGAGCAAAAAGATCCCTAAGTGGCAGAGTTATATCTACCATTAATTTCTGCAC 1586
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 962 ACACAGTAGCAGTAGAGTGAAGCGGCAATATATTTAATGATATTAATGATATGATATC 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1587 CAATGTTAGCGATGAATGCGCAACCAATGTAGTTAATTAAGTAAATATATTAATCT 1646
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1022 TATTCAGGATTCAGAGAGTCAAAATGACAGTGAATTAACGACAAGATGTT 1081
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1647 TATTTGCGGCTACCGTTTAATTCAGAGAGTAAATGATGATGATCTTGGATGAATGCTAAT 1706
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1082 AT-----ATGCTAAGGCCCGGAGCGAGCTCTTAATGAGCCAC 1120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 1707 ATGCGTTGTTGATATATGTTGCAATGTCGATATGCTGATGATGCTAATCGATCTT 1766
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1121 ACAACCCGATAATGAAGAACTGACCTTGATATGAACATGAATCTGACCTGCTGATCA 1180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1767 ATAAACCATTAATGATATTTCTGGAGTAGTCTTGACTGCTTCTGTTCTCAAACTGGCGGA 1826
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1181 CACACACTTACTCTGATTTGCGGTATCCGCAACCTG---AAGTAAATATGTTGATCTCA 1237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1827 CAGCACTTATTCATATTTATGCTGTCCCGTTGCCGGAAGATGACCAAGATTAAGTTA 1886
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1238 CAAGTTAATGACGAATAGA 1257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1887 CTTCATATATGACTAATAGA 1906
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: June 18, 2003, 15:23:44
 Job time : 211.939 secs

Om nucleic - nucleic search, using sw model

Title: US-09-986-682B-4

Scoring table: IDENTITY_NUC

Searched: 2185239 seqs, 1125999159 residues

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq Length: 20000000000

Post-processing:	Minimum match	08
------------------	---------------	----

Listing first 45 summaries

Database : N_Geneseq_101002:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1365	100.0	2408	19	AAV17621	Bacillus sp. V230
2	617.2	45.2	1449	24	ABK75353	Bacillus lichenif
3	577	42.3	10317	21	AAZ35988	Vector pgenbell s
4	575.4	42.2	1668	21	AAZ57329	SacB and cpy chime
5	575.4	42.2	1722	21	AAZ57330	SacB and cpy chime
6	542.8	39.8	1564	21	AAZ64940	DNA encoding novel
7	542.8	39.8	1935	23	AAST0417	DNA encoding novel
8	542.8	39.8	2044	23	AAZ58534	DNA encoding novel
9	542.8	39.8	2044	23	AAZ587384	DNA encoding novel

10	542.8	39.8	2335	23	AA569777	DNA encoding novel
11	542.8	39.8	2325	23	AA575941	DNA encoding novel
12	542.8	39.8	2351	12	AAQ10498	B. anthracis
13	542.8	39.8	2370	23	AA565133	DNA encoding novel
14	542.8	39.8	2370	23	AA592594	DNA encoding novel
15	542.8	39.8	2374	23	AA587392	DNA encoding novel
16	542.8	39.8	2472	23	AA589566	DNA encoding novel
17	542.8	39.8	2479	23	AA586776	DNA encoding novel
18	542.8	39.8	2562	23	AA588281	DNA encoding novel
19	542.8	39.8	2568	23	AA566887	DNA encoding novel
20	542.8	39.8	2682	23	AA571354	DNA encoding novel
21	542.8	39.8	2682	23	AA587361	DNA encoding novel
22	542.8	39.8	2742	23	AA566648	DNA encoding novel
23	542.8	39.8	2790	23	AA587386	DNA encoding novel
24	542.8	39.8	3084	23	AA585832	DNA encoding novel
25	542.8	39.8	3084	23	AA587383	DNA encoding novel
26	542.8	39.8	3267	23	AA585793	DNA encoding novel
27	542.8	39.8	3505	23	AA585838	DNA encoding novel
28	542.8	39.8	4132	23	AA586045	DNA encoding novel
29	542.8	39.8	5912	23	AA586748	DNA encoding novel
30	542.8	39.8	5938	23	AA587395	DNA encoding novel
31	542.8	39.8	6721	24	AA518599	Purificric Receptid
32	542.8	39.8	6721	24	AA518600	Purificric Receptid
33	541.2	39.6	1690	23	AA585792	DNA encoding novel
34	541.2	39.6	1690	23	AA586019	DNA encoding novel
35	541.2	39.6	1690	23	AA587350	DNA encoding novel
36	541.2	39.6	3288	23	AA577395	DNA encoding novel
37	541.2	39.6	3288	23	AA585789	DNA encoding novel
38	541.2	39.6	3288	23	AA586010	DNA encoding novel
39	541.2	39.6	3288	23	AA587346	DNA encoding novel
40	541.2	39.6	3288	23	AA588746	DNA encoding novel
41	539.6	39.5	2338	23	AA564228	DNA encoding novel
42	539.6	39.5	3305	18	AA568802	Mutant levanucras
43	532.8	39.0	1629	23	AA585800	DNA encoding novel
44	532.8	39.0	1629	23	AA586022	DNA encoding novel
45	532.8	39.0	1629	23	AA587359	DNA encoding novel

ALIGNMENTS	
RESULT 1	
AAV17621	
ID	AAV17621 standard; DNA; 2408 bp.
XX	
AC	AAV17621;
XX	
DT	04-JUN-1998 (first entry)
DE	Bacillus sp. V230 beta-fructofuranosidase genomic DNA.
XX	
XX	Beta-fructofuranosidase; xylosyl fructoside; isomaltosyl fructoside;
KM	lactosucrose; fructosyltrehalose; anticarcinogenic; growth promoter;
KW	Bifidobacteria; mineral absorption promoting activity; taste; texture;
KW	cosmetic; pharmaceutical; sweetener; ds.
XX	
OS	Bacillus sp.
XX	
XX	
FH	Location/Qualifiers
FT	1..360
FT	/*tag= a
FT	361..456
FT	/*tag= b
FT	457..1821
FT	/*tag= c
FT	/EC_number= 3.2.1.26
FT	/product= beta-fructofuranosidase
FT	/note= "Claim 7"
FT	1822..2408
FT	/*tag= d
XX	
XX	
PN	EP812915-A2.
XX	

XX Bacillus licheniformis.
OS
XX
XX WO200229113-A2.
PN
XX
XX 11-APR-2002.
PD
XX
XX 05-OCT-2001; 2001WO-US31437.
PE
XX
XX 06-OCT-2000; 2000US-0680598.
PR
XX 27-MAR-2001; 2001US-279526P.
PR
XX
XX (NOVO) NOVOZYMES BIOTECH INC.
PA
XX (NOVO) NOVOZYMES AS.
XX
XX
XX Berka R, Clausen IG;
PI
XX
XX WPI: 2002-416684/44.
DR
XX
XX
XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array -
XX
XX
XX Claim 4; SEQ ID NO 2644; 200pp; English.

The invention describes a method of monitoring differential expression of genes in a first *Bacillus* cell relative to expression of the genes in other *Bacillus* cells, comprising hybridising labelled nucleic acid probes isolated from *Bacillus* cells to a substrate containing array of *Bacillus* genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first *Bacillus* cell relative to expression of the same genes in one or more second *Bacillus* cells. The method is useful for monitoring global expression of several genes from a *Bacillus* cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which *Bacillus* cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

SQ Sequence 1449 BP; 492 A; 341 C; 313 G; 303 T; 0 other;

Query Match	45.28;	Score 617.2;	DB 24;	Length 3449;
Best Local Similarly	67.38;	Pred. No. 3.7e162;		
Matches 896; Conservative	0;	Mismatches 418;	Indels 18;	Gaps 1;

OY		12	AGACATACAGGAAGACTGTGGTTTGGCCATTCTTCACCGCGCTGCATCTGTAATAAATTC	71
Db		99	GGATTTCAGAAAGACTACGGATTTTCTCATATCCACAAGACATGACATCTCGAAAAATTC	158
OY		72	AGGACACAAAACACTCCTCATATTAAAGTGCCCAATTCATGCATGCATCAATCAAAA	131
Db		159	CGACAGCAAAGAAGACGCAATTTAAAGTTCTCCTCAATTCGATGCCAAAACATCAAAA	218
OY		132	CATTGATTCGGCAAAAGGATGATGATAGTCAGGCAACTTAATAGATTGAGATATGGA	191
Db		219	CATCCCTTCTCTCAAAAAGGGTATTAACAAAAAATGGAGCGCTGATTCGATTTGACCTAATGGA	278
OY	*	192	TAGCTGGCCACTGCAAAAAGCTGATGATGCTCGGCAATTTATCATGATATCACATCGT	251
Db		279	CAGCTGGCGGCTGCAAAATGCGCGAGGAGCGSTTGTCTACATACACAGCGCTACATCTTGT	338
OY		252	CTCCGCTTTAGCAGGTGACCACAAAACAGTAGATGATCTCCACTTCAATTATTCATCA	311

Db	339	TTTGGCGCTGGCGGCGATCCGAAAGAGCTGCATACACATCTATTTGTTTATCA	398
OY	312	AAAAGTCGGTGAATACATCGATTGACAGCTGGAATAATGCTGGAAGATATTTGAAGATAT	371
Db	399	AAAGAAAGGCGCAACTTCTATCGACAGCTGGAAAAAGCCGGCAGAGTGTTTAAAGACAG	458
OY	372	GGATTAATTTGTTCCAAATGATCCGTATCTTAAATATCAACACAGAGCTGTCAGCTTC	431
Db	459	CGACAAATTTGTTCCAGAGATCCCTACTCAAAATCATCAACACAGGAATGGCTCAGGTTTC	518
OY	432	TGCTACTTTAACCAAGATGGCGCAAGTCGGTTATTCTATACAGATTTACTCGAGTAATCC	491
Db	519	TGCCAGCTGACAAAAGACGAAAAGTCGAGCTGTTTACACAGCTTTTCCGGCACGGA	578
OY	492	TGAAGATGGTGGAAACGGGTGCTGTACCAAAATCATTTCCACTGCTCAAGTAAACTATTC	551
Db	579	ATA-----CGGACAGACAGCGTGACAAACAGCTCAGTCAATTTCTC	620
OY	552	CCAGCCGATGAGCTACACTTAAAGTCAGAGATATCTGATCATTAATCTGTCTTTGA	611
Db	621	TCAGCCGATTTGCGACAGCTCAAAATTTGACGCTAGAAATCATTAATTCGCTCTTTGA	680
OY	612	TGCGGAGACGCTACAGTTTATCAAAATTTACAGAAATTTATCGATGAAGGCAAGTGAT	671
Db	681	CGGCGCCACGGCAGGTATACCAAAACCTTGACAAATTCATTGACGAAGGAACATCACG	740
OY	672	TTACAGGTATAACCTACTTTAAGACCCCTACTATGTGAAGATAAGGCGCATTAATA	731
Db	741	CTCGCGCACAACCATACGATGAGAGACCCGATATATGGAAGAACCGCGCCATTAATA	800
OY	732	TCTGTCTTTGAAGCGAATACTGGAACCAACAGATGGTTATCAAGCGATCACTTTCAA	791
Db	801	TCTCTATTTTGAAGCCAAATACGCGAAACAAAACCGGCTACCAAGCAGACACTCCCTATT	860
OY	792	TAATAAGCTTACTATGGCGGAAGTGACCTCTTCTCCGATGTGAATAAATAAATCGT	851
Db	861	CACAGACCTCTACTACGGGGGACGAGAAGATGTTTAAAGAAAGAAAGCAGCAAGCTGT	920
OY	852	TCAAGTCTCTAAAAACAATTTGTTCTTTACCGAATGGTGCATTATGAGCATTGTGTAAT	911
Db	921	GCAAGTGCAGCAAAAGAACGCTTCGTGGCTTAAGCGGCTCTCGGAATCATCGAAT	980
OY	912	GCGCGATCACTATACAGTGAAGAAAGTGTATATGAACCATTAATGTCGATCAACACAGTAG	971
Db	981	AAATAAGCATTTACACTGAAAAAAGTCAATGAACCTTGATCGCCTCCAAATACGGTGAC	1041
OY	972	AGATTAAGTCGAAGCGCCAAATATATTAAATGAATTAATATGATCTATTCACGA	1031
Db	1041	AGATTAATTCGAAGCGCCAACTCTTCAAAATGAATGAATAATGGTATCTCTTCACGA	1101
OY	1032	TTCAAGAGGATCCAAATGAGCAGTGAATGAATTAACGCAAGAATGTTTATATGCTAG	1091
Db	1101	TTCAAGAGGATCCAAATGATCAATTTAGCGGATCGGTTCAAAAGCAATTTATATGCTGG	1161
OY	1092	GCCCGAAGCGACTCTTAATATGAGCCACACACACCGGATTAATGAACATGCACTGTAT	1151
Db	1161	CTATGATACAGGTTTCAATTAACCGGACCAATTCAGGCTTTAAACAATCCGAGCTTCT	1221
OY	1152	GACATGAATCTTGACCTGCTGATCTACACACACTTATCTCATTCGCGTATCCGCA	1211
Db	1221	GCATTTGACCGAGGATTCACAAATGACATCACTGTTACTTATTCACACTTTGCGGTACCGCA	1281
OY	1212	CCCTGAAGTATATATGTGTCTCTCAACAATATATGACAGATAGAGGCTTCTATCCACA	1271
Db	1281	GAAAAAAGCGCAGAGTGCTGATTTACACAGCTACATCAAAACAGGAGGATTTGCAACGA	1341
OY	1272	ACATCACTCTCACTGCGGGAACAGCTTGCGGTTATATATAAGGGTGTGACACATCTGG	1331
Db	1341	GCATCAAGCCAGTTTGGACCAAGCTTTTGTCTGAAGATCAAAAGGATCAAAACATCCGT	1401
OY	1332	AGGAGAAAAATAG 1343	

Db 1401 TGTCAAAACAG 1412

RESULT 3
AAZ3598/c
ID AAZ3598 standard; DNA: 10317 BP.
XX
AC AAZ3598:
XX
DT 09-FEB-2000 (first entry)
XX
DE Vector pGendell sequence SEQ ID NO:1.
XX
XX DNA sequencing vector; nested deletion; transposition; contig;
KM truncated lacZ; origin of replication; pGendell; mapping; ss.
XX
OS Synthetic.
XX
PN MO9953044-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99MO-IB00690.
XX
PR 10-APR-1998; 98US-0058746.
XX
PA (GEST) GENSET.
XX
PI Chumakov I, Tanaka H;
XX
DR WPI: 2000-023040/02.
XX
PT New DNA sequencing vectors, used for sequencing large regions of DNA
PT and mapping the location of markers -
XX
PS Claim 43; Page 93-96; 102pp; English.
XX
CC The present invention describes a vector comprising the following:
CC (a) a high copy number origin of replication (Ori) having at least one
CC cloning site, with at least one cloning site being positioned in the
CC high copy number Ori such that the ability of the high copy number Ori
CC to direct replication is not disrupted when no insert has been cloned
CC into the cloning site and (b) a low copy number Ori. Also described is a
CC truncated lacZ gene to distinguish cells with the truncated lacZ
CC gene at high copy number from cells carrying it at a low copy number.
CC where the former are dark blue when grown on medium containing Xgal and
CC IPTG and the latter are light blue when grown on medium containing Xgal
CC and IPTG. Products and methods from the present invention can be used
CC for the manipulation of DNA. They can be used for sequencing large
CC regions of DNA and mapping the locations of markers within large regions
CC of DNA. The present sequence represent a specifically claimed vector
CC
XX
XX Sequence 10317 BP: 2567 A; 2296 C; 2469 G; 2985 T; 0 other;
S0

Query Match 42.3%; Score 577; DB 21; Length 10317;
Best Local Similarity 67.2%; Pred. No. 1.5e-150;
Matches 902; Conservative 0; Mismatches 390; Indels 51; Gaps 4;

QY 1 ATGAACAGCGGGAGCTACAGAGAGACTGATGTTTGGCCATATTACAGCGGCTGACATG 60
DB 8707 ACGAACCAGGATATAGGAAACATACGCAATTCCTCCATATTACAGCGCATGATATG 8648
QY 61 CTAATAATTCAGAGACAGCAAAACAGTCTCAATTTAAAGTGCCCTCAATTAATGATCA 120
DB 8647 CTGCAAAATCCTGGAACGCAAAAATGAAAATATCAAGTCTCGAGTTGATTCGCTCC 8588
QY 121 GCAATCAAAAACATGATGATGGCAAAAGGATATGATTAAGTCAGCAACTTAATGATTTA 180
DB 8587 ACAATTAATAATATCTCTTCTGCAAAAG-----CCTG 8555
QY 181 GATGTATGGATACCTGGCCACTGCCAAAACGCTGATGTACTGGGCAAAATTAATCAGTA 240

Db 8554 GACGTTTGGGACACTGCGCATTTACAAAACGCTGACGGCACTGTGCGCAAACTATACAGGC 8495
QY 241 TATCAGATCGTCTCCGCTTTAGCAGGTGACCCAAAACAGTATGATCTCCACTTCAT 300
DB 8494 TACCACATCGCTTTGGATTAGCCGGAGATCTCTAAAATGCGGATGCACATCATCTTAC 8435
QY 301 TTATCTATCAAAAAGTCGGGTGATACATGATGATGACAGCTGGAAAATGCTGGAGAGTA 360
DB 8434 ATGTTCTATCAAAAAGTCGGGAAACTTCTATGACAGCTGGAAAACGCTGGCGGCTC 8375
QY 361 TTTGAAGATATGATTAATTTGTTCCAAATGATCCGATCTTAATATCAACACAGAG 420
DB 8374 TTTAAAGACAGCACAATTCGATGATGATGATCTTATCCGTAAGAACACACAGAA 8315
QY 421 TGGTCAGGTTCTGCTACTTTTACCAGATGGCCAGTCCGTTTATCTATACGATTAC 480
DB 8314 TGGTCAGGTTCTGACACATTTACATCTGACGGAAAATTCGTTTATCTACACTGATTTCC 8255
QY 481 TCAGTATATCCTGAAGATGGTGGAAACCGGTGCTGTAACCAAAATCATTTCAACTGCTCAA 540
DB 8254 TCCGTTAAACATTA-----CGGCAAAACAACTGACACTGACAA 8213
QY 541 GTAACCTTATCCAGCCGATGACACTTAAGTCAGTATGATGATCAATAA 600
DB 8212 GTTAACGATACAGCATCAGACAGCTCT--TTGAACATCAACGGGTATAGGATTATAAA 8156
QY 601 TCTGCTTTGATGGCGGAGACGTTACAGTTATCAAAATATTCAGCAATTATGATGAA 660
DB 8155 TCAATCTTTGA--CGGTGACGGAAAACGATCAAAATGATACAGAGTTATGATGAA 8099
QY 661 GGCAAGTGATTTGAGGATGATTAACCATCTTAAAGACCCCTCACTATGTTGAAGATTAAG 720
DB 8098 GGCAACTACAGCTGACGGCAGCAACATACGCTGAGAAATCTCTACGTAGAAAGATTA 8039
QY 721 GGCCATAATATATCTTCTTTGAAGCAATACGTGAACCAACAGATGTTATCAAGGCAT 780
DB 8038 GGCCACAATATCTTATGATTTGAAGCAAACTGAACTGAAAGATGCTACCAAGCGGAA 7979
QY 781 CAGCTTTCAATATAAGCTTACTATGCGGAGTACGCTCTTCCAGATGAAGAAA 840
DB 7978 GAATCTTTATTAATCAAAAGCTACTATGCGAAAAGCAATCAATCTTCCGCAAGAAAGT 7919
QY 841 AATTAAGCTGTTCAAGCTCAAAACAAATATCTTAAAGCAATGATGATGATGATGATG 900
DB 7918 CAAAACCTTGCAGAAAGGATTAATAAAAGCAAGCGCTAGTTAGCAAAACGGGCTCTCGGT 7859
QY 901 ATGTGATATGCGCATGACTATACAGTGAAGAGTGTATGAACCATTAATGTCATCA 960
DB 7858 ATGATTTGAGCTAAAGATGATTAACACACTGAAGAAAGTGAATGAAACCGCTATTCATCT 7799
QY 961 AATCAAGTACAGATGAAAGTGAAGCGGCAATATTTAAATGAATTAATTAATGATAT 1020
DB 7798 AACACAGTACAGATGAAAGTGAAGCGGCAATCTTTAAATGAAGAAAGCGCAATGATAT 7739
QY 1021 CTATTCACGATTAACAGATTCACAAATGACGATGATGATGATTAACGACAAAGATGTT 1080
DB 7738 CTGTTCACTGATCTCCCGGATTCAAAATATGACGATTTACGCGCATAGCTTAACGATATT 7679
QY 7678 TACATGCTGTTGTTATGTTTCTTAATCTTTAACTGGCCATACAGCGCTGAAACAAACT 7619
DB 7618 TATATGCTAGGGCCGAGGCGACTCTTAATATGAGCCACAAACCCGATTAATGAATCT 1140
QY 1141 GACTTGTATTAACATGATCTTGCACCTGCTGATCTGCACACACTTACTCTCATTTGC 1200
DB 7618 GGCCTTGTGTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7559
QY 1201 GGTATCCGCGACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 7558 GCTGTACTCAAGGAAAGAAACATATGCTGATTTACAGCTATATGACAAACAGAGGA 7499
QY 1261 TTCTATCCAGAAATCACTCTACCTGCGGACAAAGCTTGGGGTTAATTAATGAAGGCTCT 1320

Db	7498	TTCTACGCGAGCAAAACATCAACGTTTGGCGGACGCTTCGTCGTAACATCAAGCGAAG	7435
Oy	1321	GACACATCTGGAGGAGAAATAG	1343
Db	7438	AAACATCTGTTGCAAGACAG	7416
RESULT 4			
AAZ57329	AAZ57329 standard; DNA; 1668 bp.		
AAZ57329:	AAZ57329:		
03-APR-2000	(first entry)		
SacB and cpy chimeric gene #1.			
levansucrase; SacB; cpy; salt tolerance; vacuole guide peptide;			
carboxypeptidase A; chimeric gene; transgenic plant; yeast;			
Bacillus subtilis 168; Saccharomyces cerevisiae Xd;			
drought resistance; ds.			
Chimeric - Saccharomyces cerevisiae.			
Chimeric - Bacillus subtilis.			
CN1231337-A.			
13-OCT-1999.			
08-APR-1998;	98CN-0101336.		
08-APR-1998;	98CN-0101336.		
(GENE-) INST GENETICS CHINESE ACAD SCI.			
Chen S, Zhang H, Dong W;			
WPI: 2000-087902/08.			
Method for improving plant salt resistance using transgenic technology			
Claim 6; Page 2; 12pp; Chinese.			
A method has been developed for raising salt tolerance of plants using transgenic technology. The method comprises: cloning levansucrase gene SacB and vacuole guide peptide (cpy) sequence of carboxypeptidase A from Bacillus and DNA of yeast, using the two kinds of gene to create a chimeric gene, and using the chimeric gene to structure plant expression carrier, using the obtained plant expression carrier to transform a plant and screen for resistant seedlings. The method can obtain drought-resistant salt-tolerance plants, and the polymerase chain reaction (PCR) and Northern analysis of transgenic plants shows that the exogenous gene is integrated in the salt-tolerance transgenic plant. The present sequence represents a SacB and cpy chimeric gene from the present invention.			
SQ	Sequence	1668 bp; 578 A; 372 C; 325 G; 393 T; 0 other;	
Query Match	42.2%;	Score 575.4;	DB 21; Length 1668;
Best Local Similarity	67.1%;	Pred. No. 1.9e-150;	
Matches	901;	Conservative 0;	Mismatches 391; Indels 51; Gaps 4;
Oy	1	ATGAACACCGGGAGCTACAGAGACTATGCTTTGGCCATATTACACGCGTGCATG	60
Db	340	ACGAACCAAAAGCCATATAAGAAACATACGCGCATTTCCATATTACACGCGCATGATG	399
Oy	61	CTAAAAATTCGAGGACAAACAAACAGTCTCAATTAAAGTCGCTCAATTCAATGATCA	120
Db	400	CTGCAAAATCCTGGAACAGCAAAAAAATGAATAATATCAAGTTCCGATTCGATTCGTC	459
Oy	121	GCATCAAAAACATGATGATTCGGCAAAAGGGATATGATGAAGTCAAGCACTTAATGATTA	180
Db	460	ACAATTAATAATCTCTTCGCAAAAGC-----CCTG	492

QY	181	GATGATGAGGAAATAGCTGGCCACTGCAAAACCGTGTATGGCGGCAAAATATCATGGA	240
Db	493	GACGTTTGGGACAGCTGGCCATTACAAAACCGTGGCCACTGTGCACAACTATCACGGC	552
QY	241	TATCAATCACTGCTCCGCTTTAGCAGGTATACCCAAAAAACAGTATGATCTACACTTCAT	300
Db	553	TACCACTACGTCTTTGCATTTACCGGAGATCCTAAAATGGCGATGACACATCGATTTAC	612
QY	301	TTATTCATCAAAAAGTCGGTGATATCATGATTTGACAGCTGGAAAAATGCTGGAGAGTA	360
Db	613	ATGTCCTCATCAAAAAGTCGGGAAACCTCTATTGACACTGGAAAAACGCTGGCCGCTGC	672
QY	361	TTTGAAGATATGAGTAATTTTCTCCAAATATATCCGTAATCTTTAATATCAAAACACAGAG	420
Db	673	TTTAAAGACACGACAAAATTCGATGCAAAATGATTTATCTAAAAGACCAAAACACAAGAA	732
QY	421	TGTCAGAGTCTGCTACTTTTAACCAAGATGGCCAGTCCGTTATTTCTATACAGATTAC	480
Db	733	TGGTCAGGTTCAAGCCACATTTTACATCTGACCGGAAAAATCCGTTTATCTACACTGATTC	792
QY	481	TCAGGTAATCCTGAAGATGTGGAAACGGTGCCTGTAACCAAAATCTTCAACTGCTCAA	540
Db	793	TCGGTAAACATTA-----CGCAAAACAACACTGACAACTGCACAA	834
QY	541	GTAACCTATATCCACGCCGATGCACTACTTAAGTCGATGGAGTATCTGATATATAA	600
Db	835	GTTAATCGTATCAGATCAGACAGCTCT--TTGAACATCAACGGGTGAGAGATTATATAA	891
QY	601	TCCTGCTTTGAATGGCGGAGACGGTTCAGTTTATCAAAATATTTACCAATTTATCGATGAA	660
Db	892	TCAACTCTTGA---CGGTGACCGGAAAAACGTTACAAAATGTACACAGCTTCATCATGATA	948
QY	661	GGCAAGTGATTCAGTGATPAACCATCTTTAAGAGACCTCACTATGTTGAAGATAG	720
Db	949	GGCAACTCACTCAGCGGACACCATCATGCGCTGAGAGATCTCCACTACGTAGAAATATAA	1008
QY	721	GGCCATAATATCTCTCTTTGAAGCGAATACTGCAACACAGATGGTTATCAAGCCAT	780
Db	1009	GGCCAAATATCTTAGTATTTTGAACCAAACTGGAACGTGAAGTGGCTACCAACGGCAA	1068
QY	781	CAGTCTTTCAATTAATAAGCTTACTATGCGGAATGACGCTCTTCTCCAGATGAAAAA	840
Db	1069	GAATCTTATTTAACAAGCATACTATGCGCAAAACACATCATCTTCTCCGTCAAGAAAT	1128
QY	841	AATAAATGCTTCAAAAGTCTAAAAAACAAATTGCTCTTTAGCGAATGTGCTATTAGG	900
Db	1129	CAAAAACCTTCGCAACGATGATTAATAAAAGCGACGCTGATAGCAAAACGGCGCTCCGCT	1188
QY	901	ATTGTGTAATTTGGCCGATGACTATACATGTAAGAAATGTTATGAACCATTTAGTCGCATCA	960
Db	1189	ATTGATTTAGCTTAACGATGATTTACACCTGTGAAAAAAGATGTAACCCGCTGATTTGACCT	1248
QY	961	AACACAGTAGAGATGAGATCGAAGCGGCCCATATATTTAAATGAATTAATAATGATAT	1020
Db	1249	AACACAGTAGAGATGAGATTTGAAGCGCGCAACGCTTTAAATATGAACGGCAATGGTATC	1308
QY	1021	CTATTACAGGATTCACAGATCCAAAATGACGAGTGTGAATTTAACGACAAAGATGTT	1080
Db	1309	CTGTTCACATGACTCCCGCGGATCAAAAATGACGATTTGACGGCATTTACGCTTAACGATATT	1368
QY	1081	TATATGCTAAGGGCCCGAGGGGACGCTTAATGGGCCCAACAACCCGATTAATGAACCT	1140
Db	1369	TACATGCTTGGTATGATGTTCTTAATTTCTTTAACTGGCCCATATACACGCCCTGAAACAACT	1428
QY	1141	GGACTTTGATTAAGACATGAATCTTGACCCTGCTGATCTCACACACACTTTACTCTCATTC	1200
Db	1429	GGCCTTGTTAAAAAATGATCTTGATCTTAACGATGTAACTTACTTACTACACACATTC	1488
QY	1201	GGTATCCCGCACCTGGAAGGTAAATATGTTGTTACTACACAGTTATATGACAAATAGAGGC	1260
Db	1489	GCTGTAACCTCAAGGAAAGGAACAATGTCGTGATTTCAAACTATATGACAAACAGAGGA	1548

DB 1520 GCTGACCTCAAGCAAGAAACAATGCTGATTACAGCTATATGCAACAGAGCA 1579
QY 1261 TTCTATCAGAACATCACTCTCCTCGCGGACAGCTTGGGTTAATATTAAAGGCT 1320
DB 1580 TTCTACGAGACAAACATCAACCTTGGCGCAACCTTCTCTGTAACATCAAGGCAAG 1639
QY 1321 GACACATCTGGAGAGAAAATAG 1343
DB 1640 AAAACATCTGTTGTCAAGACAG 1662

RESULT 6
AAS68494
ID AAS68494 standard; cDNA; 1584 BP.
XX AAS68494;
AC
XX 13-FEB-2002 (first entry)
XX
XX
DE DNA encoding novel human diagnostic protein #4298.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HSE-) HSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI: 2001-639362/73.
DR P-PSDB: ABG04307.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX
PS Claim 1: SEQ ID No 4298; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1584 BP; 539 A; 358 C; 350 G; 337 T; 0 other;

Query Match 39.8%; Score 542.8; DB 23; Length 1584;
Best Local Similarity 65.9%; Pred. No. 2.4e-141;
Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;
QY 16 TCAAGGAAGACTATGCTTTTGGCCATTTACAGCGCTGACATGCTTAAATTCACAGA 75
DB 274 TCAAAAGAAACGTACGGCGTCCTCATATTACACGCCATGATGCTCAGATCCCTAAA 333
QY 76 CAACAACACAGTCCTCAATTTTAAGTTCAGTATGATGATGATGATGATGATGATGAT 135
DB 334 CAGCAGCAAAACAGAAATATCCAGTGCCTCATTTGATGATGATGATGATGATGATGAT 393
QY 136 GATTGGGCAAAAGGCTATGATTAAGTCAGCACTTAATAGATTGATGATGATGATGATGAT 195
DB 394 GAGTCTGCAAAAGACTT-----GATGTGGGACAGC 426
QY 196 TGGCCATGCAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
DB 427 TGGCCGCTGCAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
QY 256 GCTTTAGAGAGTGGACCAAAACAGATGATGATGATGATGATGATGATGATGATGATGAT 315
DB 487 GCTTTGGGGAGGCGCGAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 546
QY 316 GTGCGTGAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
DB 547 GTGCGGCAACACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 376 AATTTGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
DB 607 AAGTTGCAAGCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 436 ACTTAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
DB 667 ACCTTTACATCTGAGGAAATATCCGTTTATTTCTACCTGATGATGATGATGATGATGAT 725
QY 496 GATGCTGGAACCGGTGCTGATACCAATTCATTCAGTCTCAATCAATTCATTCATTCAT 555
DB 726 -----CGGCAACAAAGCCATGCAACAGCGCCAGTAAATGTCANAA 768
QY 556 CCGGATGCAAGCTACCTTAAGTGAAGTATGATGATGATGATGATGATGATGATGATGAT 615
DB 769 TCTGAT---GACACCTCAAAATCAACGAGGAGATGATGATGATGATGATGATGATGAT 822
QY 616 GGAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
DB 823 GGAGACGGAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 882
QY 676 GGTGATTAACATCTTTAAGAGACCTCATCTATGTTGAAGATTAAGGCTAATATATCTT 735
DB 883 GCGGACAAACATGCTGAGAGACCTCATCTATGTTGAAGATTAAGGCTAATATATCTT 942
QY 736 GTCTTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
DB 943 GATTTCGAAGCAACACGGAAGCAAGAAACGATACCAAGGAGGAGATCTTTATTATAC 1002
QY 796 AAGGCTTCTATGCGGGAAGTACCTCTTTCAGATGATGATGATGATGATGATGATGAT 855
DB 1003 AAGGCTTCTATGCGGGAAGTACCTCTTTCAGATGATGATGATGATGATGATGATGAT 1062
QY 856 AGTCTTAACCAATGCTTTTACCGAATGATGATGATGATGATGATGATGATGATGATGAT 915
DB 1063 AGCGTAAACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
QY 916 GATGATTAACAGTGAAGTGTATGAAACCATTAAGTGCATCAACAGTACAGAT 975
DB 1123 AATGATTAACATGTAAGAAAGTATGAAGCGGCTGATCAACAGTACAGAT 1182
QY 976 GAAGTCGAAGCGCAATATATTTAATGAAATTAATGATGATGATGATGATGATGATGAT 1035
DB 1183 GAAATCGAGCGGCAATGTTTCAAAATGAAGGCAATGATGATGATGATGATGATGATGAT 1242
QY 1036 AGAGATCAAAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095

Db 1243 CGCGGTCAAAAATGACGATGATGATTAACCTCAACAGCATTTACATGCTGTAT 1302
 1096 GGAGGCGATCTCTTAATAGCGCCACACACACCCGATTAATGAACCTGATTTAGAC 1155
 1303 GATCAAACTCTTTAAACCGCCCTTACAAAGCCGGAACAAACAGGCGCTGTGCA 1362
 1156 ATGAATCTGACCCCTGATCTACACACACTTACTCTGATGCGGATCCGACCCCT 1215
 1363 ATGGCTTGATCCAAAGATGTGACATTTCACTTCTCTCACTTGTGCAATGCCGCAACC 1422
 1216 GAAGTAAATATGTGTTACTACAAATATATGACGAATAGAGCTTCTATCCAGACAT 1275
 1423 AAAGCAACAAATGTGTTATACAAAGCTACAGCAACAGAGCTTCTGAGGATAAA 1482
 1276 CACTCTACACGCGGACAGCTGGGTTAATTAAGGTGTGACACATCTGAGAGA 1335
 1483 AAGGCAACATTTGCCCAAGCTTCTTAATGAACATCAAGGCAATTAACATCCGTTGTC 1542
 1336 GAAATAG 1343
 1543 AAAAACAG 1550

RESULT 7
 AAS70417
 ID AAS70417 standard; cDNA; 1935 BP.

AC AAS70417;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #6221.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.
 PN W0200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-0508631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB: ABG06230.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1; SEQ ID No 6221; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIR0
 CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1935 BP; 611 A; 468 C; 422 G; 434 T; 0 other;

Query Match 39.8%; Score 542.8; DB 23; Length 1935;

Best Local Similarity 65.9%; Pred. No. 2.6e-141;

Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;

QY 16 TACAAGAAAGCTATGTTTGGCCATTTATACAGCGCTGACATGTAATAAATTCACAGA 75

DB 625 TACAAGAAAGCTATGTTTGGCCATTTATACAGCGCTGACATGTAATAAATTCACAGA 684

QY 76 CAACAAACAGTCCATTTAAAGTGGCTCAATTCATGATGATGATGATGATGATGATGAT 135

DB 685 CAGCAGCAAAACGAAATATACCAAGTCCCTCAATTCATGATGATGATGATGATGATGAT 744

QY 136 GATTGCGCAAAAGGCTATGATAGTACAGCAACTTAATAGATTTAGATGATGATGATGAT 195

DB 745 GAGTGTGCAAAAGGCTATGATAGTACAGCAACTTAATAGATTTAGATGATGATGATGAT 777

QY 196 TGCCCATGCAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 255

DB 778 TGCCCATGCAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837

QY 256 GCTTTAGCAGGTGACCCCAAAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 315

DB 838 GCTTTAGCAGGTGACCCCAAAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 897

QY 316 GTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375

DB 898 GTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957

QY 376 AAATTTGTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435

DB 958 AAATTTGTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017

QY 436 ACTTTAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495

DB 1018 ACTTTAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076

QY 496 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555

DB 1077 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119

QY 556 CGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 615

DB 1120 CGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173

QY 616 GGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675

DB 1174 GGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233

QY 676 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735

DB 1234 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293

QY 736 GTCTTTGAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795

DB 1294 GTCTTTGAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353

QY 796 AAAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855

DB 1354 AAAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413

QY	856	AGTCCTAAAAAACAATTCGCTCTTTGAGCCAAATGGTGCATTAGAGCATTTGTAATGGCC	915
Db	1414	AGCCCTAAAAAACCGCATGCTGAGTTAGCAACGGCCCTCGGTATCATAGAGTTAAAT	1473
QY	916	GATGACTATACAGTGA AAAAGTGTATGA AAACCATTA GTGGCATCAACACAGTAGCAGAT	975
Db	1474	AATGATTACACATTTGAAAAAAGTAAAGCCCGCTGATTCAC TTCAAAACACGTAACGTAT	1533
QY	976	GAATTCGACGGCCGCAATATTTAATAAATGAATATAATAGTATCATTAACGAGATCA	1035
Db	1534	GAATTCGACGGCCGCGAATGTTTCAAAAATGAACGGCAAAATGTA CTGTTCACGATTCA	1593
QY	1036	AGAGGATCCAAATAGACGAGTGAATTAACGACAAAGATGTTTATATGCTAGGGCCC	1095
Db	1594	CGCGTTCAAAAATAGACGATGAGTGGTATTAACCAAAACGATTTATTCATGCTTGGTTAT	1653
QY	1096	GGAGGCGACCTCCTTAATATGGCCACACACCCGCAATTAATGAACCTGACCTGTATGAAC	1155
Db	1654	GTATCAAACTTTTAAACCGGCCCTTACAGCCGTGAACCAAAACAGAGGCTGTCTGCAA	1713
QY	1156	ATGAATCTTGACCCCTGCTGATCTCACACACACTTACTCTCATTTGCGGTATCCGCACCCCT	1215
Db	1714	ATGGGTCTTGATCCAAACGATGACATTCCTACTCTCCTTCTCGCAGT GCCGCAAGCC	1773
QY	1216	GAAGCTAATATATGGTATCTACACAAGTTATATGACGAATAGAGCTTCTATCCAGAACAT	1275
Db	1774	AAGAGCAACATATGTGTTATACACAAGCTACATGCAAAAACAAGGCTCTTCTCGAGGATAAA	1833
QY	1276	CACCTTCACCTGCGCGACAAGCTTGGGGTTAATATTAAATGAAGGTTGACACATCTGAGAGA	1335
Db	1834	AAGGCACACATTTGGCGCAAGCTTCTTAATGAAACATCAAAAGCCATTAACATCCGTTGTC	1893
QY	1336	GAATAATG 1343	
Db	1894	AAAAACAG 1901	

```

RESULT 8
AAS85834
ID AAS85834 standard; cDNA; 2044 BP.
XX
AC AAS85834;
XX
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #21638.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss
XX
OS Homo sapiens.
XX
MO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG21647.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX

```

	PS	Claim 1 SEQ ID NO 21638; 103ppr; English.
	XX	The invention relates to isolated polynucleotide (I) and
	CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
	CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
	CC	and gene mapping, and in recombinant production of (II). The
	CC	polynucleotides are also used in diagnostics as expressed sequence tags
	CC	for identifying expressed genes. (I) is useful in gene therapy techniques
	CC	to restore normal activity of (II) or to treat disease states involving
	CC	(II). (II) is useful for generating antibodies against it, detecting or
	CC	quantitating a polypeptide in tissue, as molecular weight markers and as
	CC	a food supplement. (II) and its binding partners are useful in medical
	CC	imaging of sites expressing (II). (I) and (II) are useful for treating
	CC	disorders involving aberrant protein expression or biological activity.
	CC	The polypeptide and polynucleotide sequences have applications in
	CC	diagnostics, forensics, gene mapping, identification of mutations
	CC	responsible for genetic disorders or other traits to assess biodiversity
	CC	and to produce other types of data and products dependent on DNA and
	CC	amino acid sequences. AAS64197-AAS94564 represent novel human
	CC	diagnostic coding sequences of the invention.
	CC	Note: The sequence data for this patent did not appear in the printed
	CC	specification, but was obtained in electronic format directly from WIPO
	CC	at ffp.wipo.int/pub/published_pccl_sequences.
	XX	
SQ	Sequence	2044 BP; 673 A; 507 C; 438 G; 426 T; 0 other;
	Query Match	39.8%; Score 542.8; DB 23; Length 2044;
	Best Local Similarity	65.9%; Pred No. 2.7e-141;
	Matches	875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;
OY	16	TACAGGAAGACATGTTGGTTCCTAATTACAGCGGTGCATCCTAAATAATCCAGA 75
Dd	734	TACAAAAGAACGTACGCCGCHCTCATTTATACAGCGCATGATATCTGCAGATCCCTAAA 793
OY	76	CACCAAAACAGTCCCTCAATTTAAGTGCCCTCAATTCATGCATCGACAATCAAACATT 135
Dd	794	CAGCAGCAAAACGAAAAATPACGAAGGCTCCTCAATTCGATCAATCAAGATTAAAAATTT 853
OY	136	GATTTCGGCAAAAGGATATGATTAAGTCAGGCAACTAATAGATTGATGTATGCATGAC 195
Dd	854	GAGTGTGCAAAAGGACTT-----GATGTGGGACACG 886
OY	196	TGGCCACTGCAAAACGCTGATGCTACTGCGCAAAATTATCATGATGATACATGCTCC 255
Dd	887	TGGCGCTGTCAAAACCTCACGGAACAGTAGATGATACAAAGGCTATACAGTGTGTTTT 946
OY	256	GCTTTAGCAGGTGACCACAAAAACAGTAGTATGATCCACATCTTATTCTATCAAAAA 315
Dd	947	GCTCTTTGCGGGAAGCCGGAAGAGCTGTATGACATCATCTTAATGATGTTTATCAAAAG 1006
OY	316	GTCGGTGATATCATGATTCACACTTGAAAAATGCTGGAAGATTTTGAAGATATGGAT 375
Dd	1007	GTCGGGCACAACTCATTCGACACTTGAAAAACGGGGCGTGTCTTTAAAGACAGCAT 1066
OY	376	AATTTGTTCCAATATATCCGTATCTTAATATCAACACAGAGATGGTSCAGTTTGCT 435
Dd	1067	AAGTTGACGCGCAAGATCCGATCTCGAAAGATCGACAGCAAGATGGTCCGGTTTGCA 1126
OY	436	ACTTTAAACCAAAATAGCCCAAGTCGTTTATTTCTATACGATTACTCAGTAATTCGAA 495
Dd	1127	ACCTTTAATCATCTGACGGAAAAATCCGTTATTCTACACAGACTATTCCGGTAAACATTA - 1185
OY	496	GATGTGTGAACCGGTCTGTGAACCAATATCATTTCACTGCTCAAGTAACATTAATCCCAG 555
Dd	1186	-----CGGCAAAACCAAAACCTCGAACACAGCGCAGGTAAATGATGTCAAAA 1228
OY	556	CCGATGACACTTCACCTTAAGTCGATGAGATCGATCGATTAATTCGCTTTGATGGGC 615
Dd	1229	TCTGAT---GACACACTCAAAATCAACGAGGTGGAAGATGCACAAAACATTTTGA ---C 1282
OY	616	GGAAGCGGTACACTTTACAAAATATTACGCAATTTATTCATGATGAAGCAAGCACTGGATTCA 675
Dd	1283	GGAGCGGAAAAACATATGACGAAGCTTGACGATTATTCAGATGAAGGCAATTTATCAATCC 1342

Db 918 ACCTTTACATCTGACGAAAAATCCGTTTATTCTACACTGACTATTCGCGTAACATTA- 860
 QY 496 GATGTTGACACCGGTGCTGTACCAAAATTCATTCAGTGAATTAATCTATCCAG 555
 Db 859 -----CGGCAACAAAGCCTGACACAGCGCCAGCTAAATGTCTCAAAA 817
 QY 556 CCGATGACGCTACATTAAGTCGATGAGATATCTGATCATTAATCTCTTGTATGCG 615
 Db 816 TCTGAT---GACACACTCAAAATCAAGGAGTGGAGATCAACAAACATTTTGA---C 763
 QY 616 GAGACGCTACAGTTTATCAAAATATTCAGCAATTTATGATGAAGCAAGTGGATTGCA 675
 Db 762 GAGACGCGAAAAAATATATCAAGCTTCAGCAAGTTATGATGAAGGCAATTTATACATCC 703
 QY 676 GGTGATACCATCTTATAGAGACCCCTACTATGTTAGTAAGGAGGCGCATTAATATCT 735
 Db 702 GCGCAACACATACCTGCTAGAGACCCCTACTATGTTAGTAAGGCAAGGCGCATTAATATCT 643
 QY 736 GCTTTTGAAGCGAATCTGAAACAACAGATGTTATCAAGCGATCACTCTTTCATTAAT 795
 Db 642 GTATTGCAAGCAACACGCGAACAAGAAACGATACCAAGCGAAGATCTTTATTTAAC 583
 QY 796 AAGCTTACTATGGCGGAAGTACGCTCTTCTCCAGATGAAAAAATAAATCTGCTTCA 855
 Db 582 AAAGGTTACTAGGCGGCGGACGAACTCTTCGTTAAAGAAAGCCAGAACTTCAGCAG 523
 QY 856 AGTCTTAAACAAATGCTTCTTACGAAATGGTGATTAAGGCTGTTGAATTTGGCC 915
 Db 522 AGCGCTTAAACAAACGCGATGCTGATGAGTACGAAAGCGCCCTCGTATATATAGATTAAAT 463
 QY 916 GATGATCTACAGTGAAGATGTTATGAACCATTAAGTCGATCAACACAGTACAGAT 975
 Db 462 AATGATTCACATGTAAGAAAGTATGAAGCGCGTATGATCACTTCAACACGTAACGTAT 403
 QY 976 GAAGTCGACCGCGCAATATTTTAAATGAATATTAATGATCTATTCACGATTCATCA 1035
 Db 402 GAAATCGAGCGCGGAATGTTTCAAAATGAACGCGCAATGTTACTTGTTCACGTATCA 343
 QY 1036 AGAGATCAAAATGACGAGTGAATGAATGAACGAAGATGTTATATGCTACGCGCC 1095
 Db 342 CGCGGTTCAAAATGACGATGATGATGATTAATCAACGATATTTACATGCTTGTAT 283
 QY 1096 GAGGCGACTCTTAAATGACGCGCACACACACCCGATTAATGAATGCACTTGTATGAAC 1155
 Db 282 GTATCAAACTCTTAAACGCGCCCTTAAACGCGCGTGAACAAACAGCGCTGTGCTGCA 223
 QY 1156 ATGATCTTGAACCTGCTGATCTCACACACACTTACTCTCATGCGGATATCCGACCT 1215
 Db 222 ATGGTCTTGAATCCAAACGATGTGACATTCCTCACTTCGCGAGTGGCGCAAGCC 163
 QY 1216 GAAGTATTAATGCTGATCTACAGTATATGAGCAATAGAGGCTTCTATCCAGAACAT 1275
 Db 162 AAAGCAACATGTTGTTATCATCAAGCTACATGACAAACAGAGCTTTTTCGAGATTA 103
 QY 1276 CACTCTACCTGCGGACAGCTTGGGTTAATTAAGGCTTGACACATCTGAGAGA 1335
 Db 102 AAGCAACATTTGGCGCAAGCTTCTTAATGAACATCAAAAGCAATTAACATCCGTTGTC 43
 QY 1336 GAAATATG 1343
 Db 42 AAAAACAG 35
 RESULT 10
 AAS69777 standard; cDNA: 2325 BP.
 AAS69777;
 13-FEB-2002 (first entry)
 DNA encoding novel human diagnostic protein #5581.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Dmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR P-PSDB; ABG05590.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 1; SEQ ID No 5581; 103pp: English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping; identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 SQ Sequence 2325 BP; 713 A; 593 C; 566 G; 453 T; 0 other:
 Query Match 39.8%; Score 542.8; DB 23; Length 2325;
 Best Local Similarity 65.9%; Pred. No. 2.8e-14;
 Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;
 QY 16 TACAAGGAAGACTATGCTTTGCCCCATTTACAGCGCGCTGACATGCTTAAATTCAGAGA 75
 Db 115 TACAAGGAAGACTATGCTTTGCCCCATTTACAGCGCGCTGACATGCTTAAATTCAGAGA 174
 QY 76 CAACAAACAGCTCTCATTTAAAGTGCCCTCAATTCATGACGATCAACAAACATTT 135
 Db 175 CAGCAGCAAAACGAAATATACCAAGTGGCTCAATTCATGATCAATTAACAAATATTT 234
 QY 136 GATTGCGCAAAAGGCTATGATTAAGTCAAGCACTTAATAGATTAGATGATGGAATAGC 195
 Db 235 GAGTCTGCAAAAGGACTT-----GATGTGTGGACAGC 267
 QY 196 TGGCAGTCAAAAGCGATGATGCTACGCGCAATTAATCATGATATACATGCTGCC 255
 Db 268 TGGCGGCTCAAAAGCGATGATGCTACGCGCAATTAATCATGATATACATGCTGTTT 327
 QY 256 GCTTTAGCAGGTGACCCAAACAGATGATATCTCACTTCAATTAATCTATCAAAA 315


```

Db      1384 AAAAACAG 1391
328 GCTCTTGGGGAAGCCGAAAGACGCTGATGACATCATCTTACATGTTTATCAAAAG 387
OY      316 GTCCGTGATACATGATGACAGCTGGAAAAATGCTGGAAGAGTATTGGAAGATATGAT 375
Db      388 GTCCGGCAACTCTCATGATGAGCTGGAAAAACCGGCGCTTTTAAAGACGCGAT 447
OY      376 AAATTTGTTCCAAATGATCCGTATCTTAAATATCAACACAGAGAGTGGCTGCTGCT 435
Db      448 AAGTTTCAGCGCCAGATCCGATCCTGAAAGATCAGACGAAGATGGTCCGGTTCTCA 507
OY      436 ACTTTAACCAAGATGCGCAAGTCCGTTTATTTATACAGATTACTCAGGTAATCCGAA 495
Db      508 ACCTTTACATCTGACGAAAAATCCGTTTATCTACACTGATTCGCGTAACATTTA- 566
OY      486 GATGTTGGAACCGGCTGCTGTAACCAATTCATTTCACTGCTCAGTAACTATCCAG 555
Db      567 -----CGCAAAACAAAGCCCTGACACAGCGCAGGTAATGTGCAAA 609
OY      556 CCGGATGCGAGTACACTTAAAGTCGATGAGTATGATCAATAATCTGCTTGTATGCG 615
Db      610 TCTGAT---GACACACTCAAAATCAAGAGTGGAAAGATCAACAAACGATTTTGA---C 663
OY      616 GGAGACGCTACAGTTTATCAAAATATTCAGCAATTTATGATGAAGCAAGTGATTTCA 675
Db      664 GGAGACGGAAAAACATATCAACAACGTTTACGAGTTTATGCAAGCAATATATCATCC 723
OY      676 GGTGATACCACTACTTTAGAGACCCCTACATATGTAAGTAAGGGCCATTAATCTT 735
Db      724 GCGCAACCACTACCTGAGAGACCCCTCATCTGTAAGCAAAAGGCGATTAATCTT 783
OY      736 GTCTTTGAAGCAATACTGAAACAACAGATGTTATCAAGGCGATCAGTCTTTCATAT 795
Db      784 GTATTGGAAGCAACACGGAACAGAAACGATACCAAGGAGATCTTTATTATTTAC 843
OY      796 AAAGCTTCTATGCGGGAATGACGCTCTTCTTCCAGATGAAAAAATTAACGCTTCAA 855
Db      844 AAAGGCTACTACGGGCGGCGACGAACTTCTCCCTAAAGAAAGGCAAGGCTTCAGAG 903
OY      856 AGTCTAAAAAACAATGCTTCTTTAGCGAATGTTGATTAAGGATTTGAAATGGCC 915
Db      904 AGCCGCTAAAAACGCGATGCTGATGATTACGAGCGGCCCTCGTATATAGATTAAT 963
OY      916 GATGACTATACAGTGAAGAGTGTATGAACCATTAAGTCGATCAACAGATAGACAGAT 975
Db      984 AATGCTTACATTTAAAAAGTATGAAGCGCGTATGATCTTCAACACGCTATCTGAT 1023
OY      976 GAACTGCAAGCGCCCAATATATTTAAATGAATTAATGATATCTATTCACGATTTCA 1035
Db      1024 GAAATCGAGCGCGCAATGTTTCAAAATGAAGCGCAATGTTACTTGTCTACTGATTTCA 1083
OY      1036 AGAGGATCCAAATGAGAGTGAATGATTAACGCAAAAGATGTTATATGTTAGGGCC 1095
Db      1084 CGCGGTTAAAAATACATGATGATGTTATTAATCAACGATATTTAATGCTTGGTTAT 1143
OY      1096 GGAGCGAGCTCCTTAATATGCCCACACACACCGGATTAATGAACGATGCTTATGTAAC 1155
Db      1144 GTATCAAACTCTTTAAACCGGCCCTTACAAAGCGGCTGAACAAACAGGGCTTGTGAA 1203
OY      1156 ATGAATCTTGACCTTGCTGATCTCAGACACATTAATCTCATTTGGGGATCCCGACCT 1215
Db      1204 ATGGTCTTGTATCCAAAGATGTGACATTTCACTTCTACTTCCAGTGGCGGCAAGCC 1263
OY      1216 GAAGGTATATGTGTACTCAACAGTTATATGAGATAGAGGCTTATATCCAAACAT 1275
Db      1284 AAAGCAACAAATGTGTATATCAACAGTACAGTAAACAGAGGCTTCTTGAGGATTA 1323
OY      1276 CACTCTCACTTCGCGGACACAGCTTGGGTTATATTAATTAAGGCTGACACATCTGAGA 1335
Db      1324 AAGGCAACATTTGGCCCAAGGCTTCTTATGAACATCAAAAGGCAATAAACATCCGTTG 1383
OY      1336 GAAATATG 1343
      1111 11

```

```

Db      1384 AAAAACAG 1391
RESULT 11
AA575941
ID AA575941 standard; cDNA: 2325 BP.
XX
AC AA575941;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11745.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HSE-) HXSEQ INC.
XX
PI Drmanac RT, Liu C, Yang YT;
XX
DR WPI: 2001-639362/73.
XX
DR P-PSDB; ABG11754.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11745; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 2325 BP: 713 A; 593 C; 566 G; 453 T; 0 other:

```

```

Query Match      39.8%; Score 542.8; DB 23; Length 2325;
Best Local Similarity 65.9%; Pred. No. 2.8e-141;
Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;

```

```

OY      16 TACAAGGAAGACTATGTTTGGCCCATATTCACGCGCTGACATGCAAAATTCGAGGA 75
Db      115 TACAAGGAAGACTATGTTTGGCCCATATTCACGCGCTGACATGCAAAATTCGAGGA 174

```



```

QY 76 CAACAAACAGCTCTCAATTTAAAGTCCTCAATTCATGCAATCAGCAATCAAAAACAT 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 CAGCAAGAAAACGAAAGATACCAAGTGCCTCAATTCATGCAATCAACGATTAATAAT 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 GATTGCGCAAAAGGTTATGATAGTCAGGCACTTAATAGATTAGATGTGGATAGC 195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 GAGTCTCGAAAAGGACTT-----GAGTGTGGGACAGC 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 TGCGCCATGCAAAACGCTGATGTAATGCGCAATATTCATGATATCAGATGCTGCC 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 TGCGCGGTGCAAAACGCTGAGAGCAAGTACGTGATATCAACGGCTATCAGCTGTGTT 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 GCTTACAGGTGACCAAAAACAGTGAATGATGATGATGATGATGATGATGATGAT 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 GCTCTGCGGAAACCGGAAAGACGTGATGACATCATCTCATGTTTATCAAAAG 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 GTGCGTATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 GTCGGGCAACATCAATGACAGCTGGAAGAACCGGCGCTGCTTAAAGACAGCAT 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 AAATTTGTCCAATGATCCGATCTTAATATGCAACACAGAGGTGATGATGATGAT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 AAGTTCAGCGCAACGATCCGATCTGAAGATCAGAGCAAGATGATGATGATGAT 507
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 ACTTTAAACAAAGATGCGCAAGTCCGTTATCTATACAGATTAATGATGATGAT 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 508 ACCTTACATCTGACGAAATCCGTTTATCTATACATGATGATGATGATGATGAT 566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GATGCTGGAACCGGTGCTGATCAACCAATCTTCACTGCTCAATGATGATGAT 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 -----CGCAAAACAAACGCTGACAGCGGAGTAAAGTGTCAAAA 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 CCGGTCAGCTCACTTAAAGTGAATGATGATGATGATGATGATGATGATGATGAT 615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 TCTGAT---GACACACTCAAAATCAACGAGTGAAGATCAAAAACGATTTTGA---C 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 GAGAGCGGTACAGTTTATCAAAATATTCAGCAATTCATGATGAGGCAAGTGA 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 GGAACACGAAAAACATATCAAGACGTTGAGGATTTATGATGAGGCAATTCAT 723
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 GGTGATACCTACTTAAAGGACCCCTACATGTTGAAGTAAAGGCAATATATTT 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 GGGCAGCAACATGAGGAGAGACCTCAGTACGTTGAAGCAAAAGGCAATTA 783
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 736 GCTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 GATTTGCAAGCAACAGGAGCAAGAAACGATACCAAGGCAAGATCTTTATTTAC 843
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 796 AAAGCTTACTATGCGGAGTGAAGTCTTCTCGAATGAAAAAATAAATGATGAT 855
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 AAAGCTTACTATGCGGAGTGAAGTCTTCTCGAATGAAAAAATAAATGATGAT 903
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 856 AGTCTTAAAAAACAATTTGCTTTTAGCGAATGTTGATGATGATGATGATGAT 915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 904 AGCGCTAAAAAAGCGATGATGATGATGATGATGATGATGATGATGATGAT 963
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 916 GATGATATACAGTAAAGTATGATGATGATGATGATGATGATGATGATGATGAT 975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 964 AATGATATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 976 GAACTGCAAGCGGCAATATATTTAAATGATGATGATGATGATGATGATGAT 1035
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1024 GAAATCGAGCGCGGATGTTTCAAAATGAAAGCAAAATGATGATGATGATGAT 1083
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1036 AGAGATCCAAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1084 CGCGGTTCAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1096 GAGAGCACTCTTAATGAGGCGCACACAAACCGATTAATGATGATGATGATGAT 1155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1144 GTATCAAACTTTTAACCGGCTTACAGAGCGCTGAACAAACAGGCGTGTCTGCA 1203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1156 ATGAATCTTGACCTGCTGATCTCACACACACTTACTCTCATTCGCGGTATCC 1215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1204 ATGGCTTTGATCCAAAGATGTGACATTCATCTACTCTGCGAGTCCCAAGCC 1263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1216 GAGGTATATATGCTGCTACTCACAAGTATATATGAGATTAAGGCTTCTATCAGAACAT 1275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1264 AAAGCAACATTTGCTGTTATCACAAGCTACATGACAAACAGGCTTCTGAGGATAA 1323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1276 CACTCTACCTGCGGACACACTTGGGTTATATTAAGGCTGACACATGAGAGA 1335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1324 AAGGCAACATTTGCCCAAGCTTCTTATGAACATCAAGGCAATAAACATCCGTTGTC 1383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1336 GAAATAG 1343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1384 AAAAACAG 1391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AA010498
ID AA010498 standard; DNA; 2351 BP.
XX
AC AA010498;
XX
DT 17-APR-1991 (first entry)
XX
DE B. amyloliquefaciens sacB[Bamp] fragment.
XX
KW levensucrase; sucrose regulation; Bacillus; SacB[Bamp]; ss.
XX
OS Bacillus amyloliquefaciens.
XX
FH Key Location/Qualifiers
FT sig_peptide 863..950
FT mat_peptide 951..2279
FT terminator 2291..2327
FT misc_feature 706..764
FT /tag= c
FT /note= "rho-independent"
FT /tag= d
FT /function= putative regulatory region
XX
PN WO9100913-A.
XX
PD 24-JAN-1991.
XX
PE 20-JUN-1990; 90WO-US03348.
XX
PR 07-JUL-1989; 89US-0376474.
XX
PA (DUPO ) DU POINT DE MEMOURS CO.
XX
PI Nagarajan V, Tang LB;
XX
DR MPI; 1991-051337/07.
XX
DR P-PSDB; AAR10671.
XX
PT Sucrose regulatable expression vector which can replicate -
PT derived from non-Bacillus subtilis Bacillus species for
PT expression in many bacterial species
XX
PS Example; Fig 3; 44pp; English.
XX
CC The levensucrase-encoding sequence was isolated from Bacillus
CC amyloliquefaciens lambda ZAP phage plaques. After three successive
CC screening and re-cloning steps, (using probes based on the sequence
CC of B. subtilis sacB), all the plaques were found to hybridise to the
CC probes. Clones 2A and 2C containing the putative scab[Bamp] gene
CC sequences were converted to Bluescript to give plasmids pBS300 and
CC pBS301, respectively. DNA was isolated from the plasmids and
CC digested with EcoRI in preparation for a Southern transfer. The
CC same sacB[Bsu] probes were used in the Southern hybridisation and

```

CC both hybridised only with a 1.5kb fragment. Sequence analysis of
CC this fragment showed it to be the sacB(bamP) gene. The sucrose
CC regulatory region can be incorporated into stable vectors to
CC control expression of heterologous genes in transformed B.subtilis.
CC see also AMQ10497-7.

XX Sequence 2351 BP; 739 A; 524 C; 506 G; 582 T; 0 other;

Query Match 39.8%; Score 542.8; DB 12; Length 2351;
Best Local Similarity 65.9%; Pred. No.2.9e-141;
Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;

```

OY 16 TACAAGGAAGATATGTTTGGCCATATACAGCGGTGACATCTTAAATTCACAGA 75
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 972 TACAAGGAAGATATGTTTGGCCATATACAGCGGTGACATCTTAAATTCACAGA 1031
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 76 CAACAAACAGTCCCAATTTAAAGTGGCTCAATTCATGCATCAGCAATCAAAACATT 135
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 CAGCGCAAAAGCAAAATACCAAGTGCCTCAATTGATCAATCAACGATTTAAATATT 1091
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 136 GATTGCGCAAAAGGCTATATAGTCAGGCAACTTAATAGATTATGATGATGATAGC 195
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1092 GAGTGTGCAAAAGGACTT-----GATGTGGGACAGC 1124
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 196 TGGCCACTGCAAAACGCTGATGTACTCGGCAAAATATCATGATATCATCTGCTCC 255
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 TGGCCGCTGCAAAACGCTGAGGAGACAGTATGATACAAAGGCTATCAGCTGTGTT 1184
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 256 GCTTATGAGGAGTGACCAAAACAGTATGATATCTACATCTTATTCATCAAAA 315
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1185 GCTTGTGGGAGGACCAAAAGGCTGATGATGATCAATCAATCTTATTCATCAAAAG 1244
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 316 TCGGTGATATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1245 GTCGCGCACTCACTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1304
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 376 AAATTTGTTCCAAATGATCCGATATTAAATATCAACACAGAGTGTGCTGCT 435
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1305 AAGTTGAGCGCCAAAGATCCGATCCCTGAAAGATCAGAGCAAGATGTCGCTGCTCA 1364
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 436 ACTTAAACCAAGATGAGGCAAGTCCGTTATCTATACAGATTACTCAGGTAATCCGTA 495
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1365 ACCTTAAATCTGACGGAATAATCCGTTTATTCTACCTGATATCCGTTAAACATTAA 1423
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 496 GATGTGAGACCGGTGCTGATTAACCAATTCATTCGCTCAATCAATCAATCCAG 555
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1424 -----CGGCAAAACAAAGCCTGACACAGCGCAGTAAATGTCTCAAA 1466
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 556 CCGGATGCACTGATCACTTAAAGTGCATGATGATGATGATGATGATGATGATGATG 615
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1467 TCTGAT---GACACACTCAAAATCAACGAGTGAAGATCAACAAACGATTTTGA---C 1520
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 616 GGAGCGGTGACAGTTATCAAAATATATGAGCAATTATGAGGCAAGTGAAGTTCA 675
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1521 GGAGCGGGAATAATATATGAGCACTTACAGCTTATGATGATGATGATGATGATGATG 1580
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 676 GGTGATAACCACTACTTAAAGACCCCTACTATGTTGAAGATAGGCAATTAATATCT 735
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1581 GGCGACACCACTACCTGAGAGACCCCTCCTGATGTAAGCAAAAGGCATTAATACCTT 1640
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 736 GTCTTTGAAGCAATGATGAGCAACACATGTTATCAAGGCGATCTTTCAATAT 795
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1641 GTATTCGAGCCAAACGAGGACAGAAACAGGATACCAAGGCAAGATCTTTATTAAAC 1700
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 796 AAAGCTACTATGCGGGAAGTGAAGTCTTCCAGTAAAGAAATAAATTCGTTCAA 855
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1701 AAAGCTACTACGCGCGGCGACAGACTTCTTCCTTAAGAAAGCAGAGGTTAGCAG 1760
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 856 AGTCTTAAACAAATGCTTCTTTAGCGAATGATGATGATGATGATGATGATGATGATG 915
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1761 AGCGCTTAAACAAACGCGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1820
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 916 GATGACTATACAGTGAAGAGTGTATGAAACCAATTAGTCGATCAACACAGTACAGAT 975
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1821 AATGATTACACATGTAAGAAAAAGTATAGAGCCGCTGATCACTTCAACACGTAATCAT 1880
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 976 GAAGCGCAAGCGGCAATATATTTAAATAATGATGATGATGATGATGATGATGATGAT 1035
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1881 GAAATCGAGCGCGGAATGTTTCAAAATGAAACGCAATGATGATGATGATGATGATGAT 1940
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1036 AGAGATCCAAATAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1941 CCGGTTTCAAAATAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2000
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1096 GGAGCGCACTCTTAATATGSCCAACACACACCCGATTAATGATGATGATGATGATGAT 1155
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2001 GTATCAAACTCTTTAAACCGCGCTTACAGCGCTGAAACAAACAGGCGTTGTCTGCA 2060
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1156 ATGATCTTGACCCCTCGCATCTGACACACACTACTCTGCTTGGGATATCCGACCCCT 1215
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2061 ATGGCTTGTATCCAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 2120
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1216 GAAGTATATATGTTGTTACTCACAAGTATATGAGATAGAGCTTCTATCCAGAACAT 1275
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2121 AAAGCAACATATGTTGTTATCACAAGTATGATGATGATGATGATGATGATGATGAT 2180
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1276 CACTCTACCTGCGGAGCAAGCTTGGGTTAATTAAGGCTGACACATCTGAGAGA 1335
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2181 AAGGCACATTTGGGCCAAGCTCTTAAATGAACATCAAAAGCAATAAACATCCGTTGTC 2240
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1336 GAAATATG 1343
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2241 AAAACAG 2248
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

AAS69133
ID AAS69133 standard; cDNA: 2370 BP.

AAS69133;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #4937.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

W0200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI: 2001-639362/73.

P-PSDB: ABG04946.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.

Claim 1; SEQ ID No 4937; 103bp; English.

The invention relates to isolated polynucleotide (I) and
polypeptide (II), sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS6197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPRO
 CC at ftp.wipro.int/pub/published_pct_sequences.

Sequence 2370 BP: 743 A; 535 C; 544 G; 548 T; 0 other:

Query Match 39.8%; Score 542.8; DB 23; Length 2370;
 Best Local Similarity 65.9%; Pred. No. 2.9e-141;
 Matches 873; Conservative 0; Mismatches 402; Indels 51; Gaps 4;

```

OY 16 TACAAGAGAGACTGTTTGGCCATTTACAGCGCGCTGACATGCTAAATTCAGGA 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 TACAAGAGAGAGCTAGCGGCTCTCATATTTACAGCCCTGATTCGTCAGATCCCTAAA 1119
OY 76 CACAAAGACGCTCTCAATTTAAAGTGGCTCAATTCAGTCAGCAATCAAAACAT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 CAGAGCAAAAGCAAAATACCAAGTGCCTCAATTCGATCAACGATTAATAATAT 1179
OY 136 GATTGCGCAAAAGGATGATGATAGTACAGCAACTTAATAGATTAGATGATGGATGC 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1180 GAGTCTGCAAAAGACTT-----GATGTGGGACAGC 1212
OY 196 TGGCCACTCAAAACGCTGATGCTACGCGCAATTAATCATGATACATGCTCTCC 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1213 TGGCGCTGCAAAACGCTGAGCAAGAGTACTGATTCACACGGGTATCAGGTGTGTT 1272
OY 256 GCTTACAGAGTGACCCCAAAACAGTATGATCTCCACTTCATTATCTATCAAAA 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1273 GCTGTGGGGAACCCGAAAGCGCTGATCACAATCATCTGATTTATCAAAAG 1332
OY 316 GTCGATGATACGATTGACAGCTGGAATAATGCGAAGATTTGAAGATATGAT 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1333 GTCGCGCAACACTCAATCGACAGCTGGAATAACCGGGCGCTGCTTTAAAGACAGCGAT 1392
OY 376 AAATTTGTTCCAAATGATCCGATCTTAAATATCAAAACAGAGAGTGTGAGTTCTGCT 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1393 AAGTTGAGCGCAACGATCCGATCTGAAGATCAGAGCAAGATGTGCGGTTCTGCA 1452
OY 436 ACTTTAACCAAGATGCGCAAGTCCGTTTATTCTATACAGATTACTCAGTAATCTGAA 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1453 ACCTTTACATCTGCGGAAATCCGTTTATTCTATACATGACTATTCGCGTAACATTA- 1511
OY 496 GATGATGGAACCGGTGCTGTAACCAATCATTTCACTGCTCAAGTAACCTATCCAG 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1512 -----CGGCAACCAAGCTGACACAGCGCGAGTAATAAGTGTCAAAA 1554
OY 556 CCGGATGACAGTACACTTAAAGTCAGAGATATCTGATTAATGCTCTTGATGCG 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1555 TCTGAT---GACACACTCAAAATCAAGGAGTGAAGATCACAAAAGCATTTTGA---C 1608
OY 616 GGAACGCTACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGGCAAGTGGATTCA 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1609 GGAACGCAAAACATATCAACAGTTTCAGCATTTAAGGCAAGTAATTAACATCC 1668
OY 676 GGTATTAACCATCTTATAGAGACCTTCACATGATGTAAGATGAAGGCATTAATATCT 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1669 GGCACAAACATACGCTGAGAGACCTCTACAGTTGAAGCAAAAGGCCATTAATACCT 1728

```

```

OY 736 GCTTTGAAGCAAGTACTGGAACAACAGATGTTATCAAGCGCATCTTCAATAT 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1729 GTATTGCAAGCAACAGCGGAACAGAAACGATTAACAGGCAAGATCTTTATTATAC 1788
OY 796 AAAGCTTCTATGCGGGAAGTGAAGCTCTTCCAGAAATGAATAAATTAACCTTCA 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1789 AAAGGCTACTACGGCGGCGGACAGAACTTCTCCGTAAGAAAGCCAGACCTTCAGAG 1848
OY 856 AGTCTAAACAAATGATCTTCTTACGGAATGGTGATAGCATGTTGAATGGCC 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1849 AGCGTTAAACACGCGATGCTGAGTTAGCGAAGCGCGCTCTGGATATCATAGATTAAAT 1908
OY 916 GATGACTATPACAGTAAGTAAGTGTATGAACCATTAAGTCGATCAACACAGTAGCAGAT 975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1909 AATGATTTACATTTAAATAAAGTATGAAGCCGCTGATCATTCAACACGTAATCTAT 1968
OY 976 GAAGTCGACCGCGCAATATATTAATAATGAATAAATGAATGATCTATTCACGGAATCA 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1969 GAAATCGAGCGCGCAATGTTTCAAAATGAACGCAAAATGATGTTCTCATCTGATTC 2028
OY 1036 AGAGATCCAAATGACGAGTATGATGATTAACGCAAGATGTTATATGCTAGGCGCC 1095
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2029 CGCGGTTAAATAATGACGATGATGATTAACATCAACGATTAATACATCTGTTAT 2088
OY 1096 GAGGCGACTCTTAAATGACCCACACACACCGATTAATGAACGATGATGATGATGAT 1155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2089 GTATCAAACTCTTTATACCGCGCTTACAAAGCGCTGAACAAACAGGCTTGTCTGCA 2148
OY 1156 ATGAATCTTGAACCTGCTGATCTCAGACACACTTACTCTCATTCGCGTATCCGACCT 1215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2149 ATGGTCTTGTATCCAAAGATGATGATCACTTCACTTCTCACTTCTGCGAGTCCGAC 2208
OY 1216 GAAGTATATATGCTGCTACTCAGCAATATATGAGATTAAGGCTTATCTATCCAGAT 1275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2209 AAAGCAACATATGCTTATCAGACTTATGACATGACAAACAGAGGCTTTTGAGGATTA 2268
OY 1276 CACTCTCACTGCGGAGCAAGCTTGGGTTATATTAAGGCTGACACATCTGAGAGA 1335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2269 AAGCAACATTTGCGCCCAAGCTTCTTATGAACATCAAGCAATTAACATCGTGTGC 2328
OY 1336 GAAATAG 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2329 AAAAACAG 2336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
AAS92594
ID AAS92594 standard; cDNA: 2370 BP.
XX
XX AAS92594;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
DE DNA encoding novel human diagnostic protein #28398.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX

```

DR WPI: 2001-639362/73.
DR P-PSDB: ABG28407.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS Claim 1: SEQ ID No 28398; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 2370 BP; 743 A; 535 C; 544 G; 548 T; 0 other;
Query Match 39 8%; Score 542.8; DB 23; Length 2370;
Best Local Similarity 65.9%; Pred. No. 2.9e-141;
Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;
OY 16 TACAAGAGAGATAGTTTGGCCATATTACAGCGGCGATGTAATAATCCAGGA 75
DB 1060 TACAAGAGAGATAGTTTGGCCATATTACAGCGGCGATGTAATAATCCAGGA 1119
OY 76 CAACAAGAGATAGTTTGGCCATATTACAGCGGCGATGTAATAATCCAGGA 135
DB 1120 CAACAAGAGATAGTTTGGCCATATTACAGCGGCGATGTAATAATCCAGGA 1179
OY 136 GATGGGCAAAAGGATGATAGTACAGCAACTTATGATTTGATGATGATGATG 195
DB 1180 GATGGGCAAAAGGATGATAGTACAGCAACTTATGATTTGATGATGATGATG 1212
OY 196 TGGCCACTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 255
DB 1213 TGGCCACTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 1272
OY 256 GCTTTAGCAGGTGACCCCAAAAGGATGATGATGATGATGATGATGATGATG 315
DB 1273 GCTTTAGCAGGTGACCCCAAAAGGATGATGATGATGATGATGATGATGATG 1332
OY 316 GTCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
DB 1333 GTCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1392
OY 376 AATTTGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
DB 1393 AATTTGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1452
OY 436 ACTTTAACCAGATGCGGATGATGATGATGATGATGATGATGATGATGATGATG 495
DB 1453 ACTTTAACCAGATGCGGATGATGATGATGATGATGATGATGATGATGATGATG 1511
OY 496 GATGGGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 555
DB 1512 GATGGGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1554

OY 556 CCGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
DB 1555 TCTGAT----GACACACTCAAAATGACAGAGAGAGAGAGAGAGAGAGAGAG 1608
OY 616 GGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 675
DB 1609 GGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1668
OY 676 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
DB 1669 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1728
OY 736 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
DB 1729 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
OY 796 AAGGCTTACTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 855
DB 1789 AAGGCTTACTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1848
OY 856 AGTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
DB 1849 AGTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1908
OY 916 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
DB 1909 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1968
OY 976 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1035
DB 1969 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028
OY 1036 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 2029 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
OY 1096 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155
DB 2089 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2148
OY 1156 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
DB 2149 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2208
OY 1216 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
DB 2209 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2268
OY 1276 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1335
DB 2269 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2328
OY 1336 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
DB 2329 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2388
RESULT 15
AAS87392
ID AAS87392 standard; cDNA; 2374 BP.
XX AAS87392;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #23196.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX

PN WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
DR P-PSDB: ABG23205.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 1: SEQ ID No 23196; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 2374 BP; 898 A; 522 C; 490 G; 464 T; 0 other:
XX
XX Query Match 39.8%; Score 542.8; DB 23; Length 2374;
XX Best Local Similarity 65.9%; Pred. No. 2.9e-141;
XX Matches 875; Conservative 0; Mismatches 402; Indels 51; Gaps 4;
XX
OY 16 TACAAGGAGACTATGTTTGGCCATATTTACGCGGTGACGTAAATTCACAGA 75
DB 1064 TACAAGGAGACTATGTTTGGCCATATTTACGCGGTGACGTAAATTCACAGA 1123
OY 76 CACAAACAGTCTCAATTTAAAGTGCCTCAATTCATGATGATGATGATGATGAT 135
DB 1124 CAGCAGCAAAACGAAATATACCAAGTGCCTCAATTCATGATGATGATGATGAT 1183
OY 136 GATTGGCAAAAGGCTATGATAGTACGCAACTTAATAGATTAGATGATGATGATGAT 195
DB 1184 GAGTCTGCAAAAGGACTT-----GATGTGGGAGACGC 1216
OY 196 TGGCGACGTGCAAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 255
DB 1217 TGGCGCGTGCAGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276
OY 256 GCTTTAGCAGGTGACCAAAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 315
DB 1277 GCTTTAGCAGGTGACCAAAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 1336
OY 316 GTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375

DB 1337 GTCGCGCAACATCAATCGACAGCTGCAAAAAAGCGGGCGGTCTTTAAAGACGCGAT 1396
OY 376 AAATTTGTTCCAAATGATCCGTATCTTAAATATCAACACAGAGAGTGTGCTGCT 435
DB 1397 AAGTTCGACGGCAACGATCCATCCTGAAGATAGAGCAAGATGCTCGGTTCTGCA 1456
OY 436 ACTTTAACCAAGAGTGGCAAGTCTTATCTATACAGATTTCTCAGTATTCACGAA 495
DB 1457 ACCTTTACATCTGACGAGAAATATCCGTTATTTCTACACTGACTATTCGCGTAAACATTA 1515
OY 496 GATGGTGAACCGGTGCTGTATACCAAAATTCATTCACCTGCTCAAGTAACCTATCCAG 555
DB 1516 -----CGGCACAAACAGCCCTGACAAACAGCGGCAAGTAAATGTGCAAA 1558
OY 556 CCGGATGACGTACACTTAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 615
DB 1559 TCTGAT---GACACACTCAAAATCAACGAGTGAAGATCAAAACGATTTTGA---C 1612
OY 616 GGAGACGGTACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGCAGTGGATTCA 675
DB 1613 GGAGACGGAAAAACATATACAGACGTTTACGACATTTATCGATGAAGCAATATATACATCC 1672
OY 676 GGTGATTAACATACCTTTAAGAGACCCCTCATATGTTGAAGTAAGGGCCATAAATATCTT 735
DB 1673 GCGCAGCAACCATACGCTGAGAGACCCCTCATACGTTGAAGCAAAAGCCATAAATATCTT 1732
OY 736 GTCCTTTGAAGGATCTGAGAACACAGATGTTATCAAGCGATCACTTTCAATAT 795
DB 1733 GTATTGCAAGCCACACCGGACAGAAACAGAAATCAAGGAGAAATCTTTATTTAAAC 1792
OY 796 AAAGCTTACTATGCGGAGAGTACGCTTCTTCCAGATGAAAAAATAAATCACTCTTCA 855
DB 1793 AAAGGCTACTAGCGGGGCGGACAGACACTTCTCCGTAAGAAACGACACACTTCACAG 1852
OY 856 AGCTCTAAAAACAATTCCTTTTACGCAATGCTGATTAAGCATTTGGAATGCGC 915
DB 1853 AGCGTAAAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1912
OY 916 GATGCTATTAACAGTAAAGTGTATGAACCATTTAGTGGATCAACACAGTACGAT 975
DB 1913 AATGTTTACATTTGAAAGAAAGTAAAGTAAAGCCGCTGATCAGCTTCAACACGTAACGTAT 1972
OY 976 GAAGTCGAAACGCGCAATATATTTAAATGAATTAATATGATGATGATGATGATGATGAT 1035
DB 1973 GAATTCAGCGCGCGGATGTTTCAAAATGAACGCAAAATGCTACTTCTTACATGATTTCA 2032
OY 1036 AGAGATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 2033 CGCGGTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2092
OY 1096 GGAGGCGACCTCTTAATGAGCCGACACACCCGATTAATGAACCTGACCTGTATGAC 1155
DB 2093 GTATCAAACTCTTTTAAACCGCCCTTACAAAGCGCTGAAACAAAGAGGCTGTGCTGCA 2152
OY 1156 ATGATCTTGAACCTGCTGATCTCAACACACTTACTCTTGGGATATCCGACACCT 1215
DB 2153 ATGGGCTGTGATCCAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2212
OY 1216 GAGGTAATTAATGTTGATCTACAGATTAATGAGCAATAGAGGCTTCTATCGAATAT 1275
DB 2213 AAAGCAACAAATGTTGTTATCAAGCTATGACAAACAGAGCTTCTTGAAGATATA 2272
OY 1276 CACTCTCACTGCGGAGCAACCTTGGGTTTAAATTAAGGCTGACACATCTGAGAGA 1335
DB 2273 AAGGCAACATTTGGCGCAAGCTTCTTAATGAACATCAAAAGCAATCAAAACATCGTTGTC 2332
OY 1336 GAAATAG 1343
DB 2333 AAAAAGAG 2340

Search completed: June 18, 2003, 03:10:28
Job time : 307.364 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:52:36 ; Search time 63.3117 Seconds
(without alignments)
6611.951 Million cell updates/sec

Title: US-09-986-682B-4

Perfect score: 1365

Sequence: 1 ATGAACAGCGGGGAGCTACAA.....CCGACACAGACATTCACA 1365

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1365	100.0	1365	2	US-08-870-827-4
2	1365	100.0	1365	4	US-09-317-179-4
3	1365	100.0	2408	4	US-08-870-827-5
4	1365	100.0	2408	4	US-09-317-179-5
5	577	42.3	10317	3	US-09-058-746-1
6	577	42.3	10317	3	US-09-438-142-1
7	539.6	39.5	3305	3	US-09-068-043-1
8	43.4	3.2	7218	1	US-08-232-463-14
9	39.8	2.9	1662	1	US-08-671-947-1
10	36.6	2.7	7101	1	US-08-480-604-9
11	36.6	2.7	7101	2	US-08-405-496A-9
12	36.6	2.7	7101	4	US-08-915-136-9
13	36.6	2.7	7101	4	US-08-857-310-9
14	36.6	2.7	3095	6	5231168-1
15	35.4	2.6	849	4	US-09-648-520E-48
16	35.4	2.6	3317	4	US-09-193-562D-1
17	34.8	2.5	62	3	US-09-068-043-5
18	34.2	2.5	3360	1	US-07-712-833A-1
19	34	2.5	1467	4	US-09-134-001C-1409
20	34	2.5	2817	5	PCT-US93-05944-1
21	33.8	2.5	3356	4	US-09-379-523-4
22	33.6	2.5	3279	4	US-08-446-137B-1
23	33.6	2.5	2408	2	US-08-870-827-5
24	33.6	2.5	2408	4	US-09-317-179-5
25	33.6	2.5	3713	1	US-08-100-709-1
26	33.6	2.5	3713	1	US-08-476-865-1
27	33.6	2.5	3713	1	US-08-474-038-1

28	33.6	2.5	3713	2	US-08-779-046-1	Sequence 1, Appl
29	33.6	2.5	3713	2	US-08-881-340-1	Sequence 1, Appl
30	33.4	2.4	1566	4	US-09-134-001C-2802	Sequence 2802, Ap
31	33.2	2.4	900	4	US-08-879-098-1	Sequence 1, Appl
32	33.2	2.4	900	4	US-09-631-548-1	Sequence 1, Appl
33	33.2	2.4	3891	1	US-08-480-604A-27	Sequence 27, Appl
34	33.2	2.4	3891	2	US-08-405-496A-27	Sequence 27, Appl
35	33.2	2.4	3891	4	US-08-915-136-27	Sequence 27, Appl
36	33	2.4	699	4	US-08-936-165A-233	Sequence 233, App
37	33	2.4	2047	4	US-08-836-261A-1	Sequence 1, Appl
38	33	2.4	4163	4	US-09-004-838-70	Sequence 70, Appl
39	33	2.4	5820	4	US-09-029-213B-7	Sequence 1641, Ap
40	32.8	2.4	630	4	US-09-134-001C-1641	Sequence 1, Appl
41	32.8	2.4	2600	3	US-08-988-251-1	Sequence 1, Appl
42	32.8	2.4	2600	3	US-09-386-048-1	Sequence 1, Appl
43	32.8	2.4	3280	1	US-08-259-000-4	Sequence 4, Appl
44	32.8	2.4	3280	1	US-08-729-767-6	Sequence 6, Appl
45	32.8	2.4	8195	4	US-08-961-527-94	Sequence 94, Appl

ALIGNMENTS

```
RESULT 1
US-08-870-827-4
: Sequence 4, Application US/08870827
: Patent No. 5962297
: GENERAL INFORMATION:
: APPLICANT: Tsusaki et al.
: TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE
: TITLE OF INVENTION: ACTIVITY
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/870,827
: FILING DATE: 06-JUN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 170,630/1996
: FILING DATE: 10-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWDY, Roger L.
: REGISTRATION NUMBER: 25,618
: REFERENCE/DOCKET NUMBER: TSUSAKI=2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1365 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-870-827-4
:
: Query Match 100.0%; Score 1365; DB 2; Length 1365;
: Best Local Similarity 100.0%; Pred. No. 0;
: Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 1 ATGAACAGCGGGGAGCTACAAAGAACTATGTTTGGCCCATATTTACAGCGCTGACATG 60
: Db 1 ATGAACAGCGGGGAGCTACAAAGAACTATGTTTGGCCCATATTTACAGCGCTGACATG 60
```

OY	61	CTAAATAATTCGAGACACAAACAGCTGCTCAATTTAAAGTCCCTCAATGCATGCATCA	120
Db	61	CTAAATAATTCGAGACACAAACAGCTGCTCAATTTAAAGTCCCTCAATGCATGCATCA	120
OY	121	GCAATCAAAAACATTGATTGGCGAAAGGGATATGATTAAGTCAGGCACCTTAATAGATTYA	180
Db	121	GCAATCAAAAACATTGATTGGCGAAAGGGATATGATTAAGTCAGGCACCTTAATAGATTYA	180
OY	181	GATGTATGGGATAGCTGGCCACTGCAAAAACGTGATGTAATGCTGGGCAATTTATATGGA	240
Db	181	GATGTATGGGATAGCTGGCCACTGCAAAAACGTGATGTAATGCTGGGCAATTTATATGGA	240
OY	241	TATCAACATGCTGCTCCGCTTTAGCAGTGAACCCAAAAACAGTATGATCTCCACTTCAT	300
Db	241	TATCAACATGCTGCTCCGCTTTAGCAGTGAACCCAAAAACAGTATGATCTCCACTTCAT	300
OY	301	TTATTTCTATCAAAAAGTCGGTGATACATCGATTGACGTGGAAAAATCGTGGAGAGTA	360
Db	301	TTATTTCTATCAAAAAGTCGGTGATACATCGATTGACGTGGAAAAATCGTGGAGAGTA	360
OY	361	TTTGAAGATATGATTAATTTGTCCAATGATGCTGATCTTAATATCAAAACACAGGAG	420
Db	361	TTTGAAGATATGATTAATTTGTCCAATGATGCTGATCTTAATATCAAAACACAGGAG	420
OY	421	TGGTCAGGTTCTGCTACTTTTAAACCAAGATGGCCAAAGTCGGTTTATCTATACAGATTAC	480
Db	421	TGGTCAGGTTCTGCTACTTTTAAACCAAGATGGCCAAAGTCGGTTTATCTATACAGATTAC	480
OY	481	TCAGGTATTCGGAAGATGTGGGAACCGGCTGCTGTAACCAATCATTTCAACCTCTCA	540
Db	481	TCAGGTATTCGGAAGATGTGGGAACCGGCTGCTGTAACCAATCATTTCAACCTCTCA	540
OY	541	GTAAACTTATCCACACCGGATGACACCTACTTAAGTCGATGAGTATCTGATCATTAA	600
Db	541	GTAAACTTATCCACACCGGATGACACCTACTTAAGTCGATGAGTATCTGATCATTAA	600
OY	601	TCTGCTTTGATGGCGGAGACGGTACAGTTTATCAAAATATTACACATTTATGCGATGA	660
Db	601	TCTGCTTTGATGGCGGAGACGGTACAGTTTATCAAAATATTACACATTTATGCGATGA	660
OY	661	GGAAGTGCATTTACGATTAACCATCTTAAAGACCCCTCATGTTGTAAGTATAG	720
Db	661	GGAAGTGCATTTACGATTAACCATCTTAAAGACCCCTCATGTTGTAAGTATAG	720
OY	721	GGCCTTAATATCTTGTCTTTGAGCGAATCTGGAACACAGATGTTATCAAGCGCAT	780
Db	721	GGCCTTAATATCTTGTCTTTGAGCGAATCTGGAACACAGATGTTATCAAGCGCAT	780
OY	781	CAGTCTTCAATTAATTAAGCTTACTATGAGCGGAATGACGCTTCTTCACAGATGA	840
Db	781	CAGTCTTCAATTAATTAAGCTTACTATGAGCGGAATGACGCTTCTTCACAGATGA	840
OY	841	AATAAATGCTTCAAAAGTCTTAAAAAACAAATTCGCTTTTAAAGCAATGGTGCAATTAAGC	900
Db	841	AATAAATGCTTCAAAAGTCTTAAAAAACAAATTCGCTTTTAAAGCAATGGTGCAATTAAGC	900
OY	901	ATTGTTGAATTTGGCCGATGACTATACATGTGAAGAACTGTTATGAAACCATTTAGTCCATCA	960
Db	901	ATTGTTGAATTTGGCCGATGACTATACATGTGAAGAACTGTTATGAAACCATTTAGTCCATCA	960
OY	961	AACACAGTAGAGATGAGTGAACGGCCCATATATTTAAATTAATTAATTAATGAT	1020
Db	961	AACACAGTAGAGATGAGTGAACGGCCCATATATTTAAATTAATTAATTAATGAT	1020
OY	1021	CTATTACGGAATCAAGAGATCCAAAATGAACGAGTGAATTAACGAACAAAGATGTT	1080
Db	1021	CTATTACGGAATCAAGAGATCCAAAATGAACGAGTGAATTAACGAACAAAGATGTT	1080
OY	1081	TATATGCTAGGAGGCGGAGGAGCTCCCTTAATGAGCCACAAACCCGATTAATGA	1140
Db	1081	TATATGCTAGGAGGCGGAGGAGCTCCCTTAATGAGCCACAAACCCGATTAATGA	1140

Accession	Sequence	Length
QY	1141 GGACCTGTGATTGAACATCAATCTTGACCCCTGCTATCTCAACACACACTTACTCTCATTTGC	1200
Db	1141 GGACTGTGATTGAACATCAATCTTGACCCCTCTGTCTCAACACACTTACTCTCATTTGC	1200
QY	1201 GGTATCCGCGACCCCTGGAAGTAAATATGTGTACTCAACAGTTATATGACGAATAGAGGC	1260
Db	1201 GGTATCCGCGACCCCTGGAAGTAAATATGTGTACTCAACAGTTATATGACGAATAGAGGC	1260
QY	1261 TTCTATCCAGAACATCACTCTCACCTGCGGGACAAGCTTGGGGTTAATATTAAAGGCTCT	1320
Db	1261 TTCTATCCAGAACATCACTCTCACCTGCGGGACAAGCTTGGGGTTAATATTAAAGGCTCT	1320
QY	1321 GACACATCTGGAGGAGAAATAGTTCGCGACGAAGAACAATTTCCCA	1365
Db	1321 GACACATCTGGAGGAGAAATAGTTCGCGACGAAGAACAATTTCCCA	1365

```

RESULT 2
US-09-317-179-4
; Sequence 4, Application US/09317179
; Patent No. 6383769
GENERAL INFORMATION:
APPLICANT: Tsusaki et al.
TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE ACTIVITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/317,179
FILING DATE: 24-May-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/870,827
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1365 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-317-179-4
Query Match          100.0%; Score 1365; DB 4; Length 1365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATGAACAGCGGGGACTACAGAGAACTATGTCTTTGCCCATATTTACACGGCGTGACATG 60
DB      1 ATGAACAGCGGGGACTACAGAGAACTATGTCTTTGCCCATATTTACACGGCGTGACATG 60
QY      61 CTAATAAATTCAGAGCAACAAGAAGAGTCCATTTAAAGGCCTCAATTCATGACATCA 120
DB      61 CTAATAAATTCAGAGCAACAAGAAGAGTCCATTTAAAGGCCTCAATTCATGACATCA 120

```


121 GCATCAAAACATGATTCGGCAAAAGGATATGATGATGACGCACTTAATAGATT 180
121 GCAATCAAAACATGATTCGGCAAAAGGATATGATGATGATGATGATGATGAT 180
121 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
181 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
181 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
241 TATCATCATGCTCTCGCTTTAGCAGGTGACCAAAACAGTATGATGATGATGAT 300
241 TATCATCATGCTCTCGCTTTAGCAGGTGACCAAAACAGTATGATGATGATGAT 300
301 TTAATCTATCAAAAGTGGTATGATGATGATGATGATGATGATGATGATGAT 360
301 TTAATCTATCAAAAGTGGTATGATGATGATGATGATGATGATGATGATGAT 360
361 TTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
361 TTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
421 TGGTCAGGTTCTGCTACTTAAACCAAAAGTGGTATGATGATGATGATGATGAT 480
421 TGGTCAGGTTCTGCTACTTAAACCAAAAGTGGTATGATGATGATGATGATGAT 480
481 TCAGGTAATCTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
481 TCAGGTAATCTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
541 GTAACTATATCCGACCGGATGATGATGATGATGATGATGATGATGATGATGAT 600
541 GTAACTATATCCGACCGGATGATGATGATGATGATGATGATGATGATGATGAT 600
601 TCTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
601 TCTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
661 GCGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
661 GCGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 GGCATTAATATCTGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 GGCATTAATATCTGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 CAGCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
781 CAGCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
841 AATAAAGTCTTCAAAAGTCTTCAAAAGTCTTCAAAAGTCTTCAAAAGTCTTCAAA 900
841 AATAAAGTCTTCAAAAGTCTTCAAAAGTCTTCAAAAGTCTTCAAAAGTCTTCAAA 900
901 ATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
901 ATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
961 AACACAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961 AACACAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021 CTATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021 CTATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1081 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1081 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1141 GGACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1141 GGACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1201 GGTATCCCGACCTGGAAGTATATGATGATGATGATGATGATGATGATGATGATGAT 1260

1201 GGTATCCCGACCTGGAAGTATATGATGATGATGATGATGATGATGATGATGATGAT 1260
1261 TTTATCCAGAACATCACTCTCAGCTGCGGACCAAGCTTGGGCTTAATTAAGGCTCT 1320
1261 TTTATCCAGAACATCACTCTCAGCTGCGGACCAAGCTTGGGCTTAATTAAGGCTCT 1320
1321 GACACATCTGGAGGAGAAATAGTTCGAGACAAAGGACAAATTCCTCA 1365
1321 GACACATCTGGAGGAGAAATAGTTCGAGACAAAGGACAAATTCCTCA 1365

RESULT 3
US-08-870-827-5
Sequence 5, Application us/08870827
Patent No. 5962297

GENERAL INFORMATION:
APPLICANT: Tsusaki et al.
TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/870,827
FILING DATE: 06-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 170,630/1996
FILING DATE: 10-JUN-1996

ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-737-3528
TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2408 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus sp.
INDIVIDUAL ISOLATE: V230 (FERM BP-5054)

FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..360
IDENTIFICATION METHOD: E
NAME/KEY: signal peptide
LOCATION: 361..456
IDENTIFICATION METHOD: S
NAME/KEY: mat peptide
LOCATION: 457..1821
IDENTIFICATION METHOD: S
NAME/KEY: 3'UTR
LOCATION: 1822..2408
IDENTIFICATION METHOD: E

US-08-870-827-5
Query Match 100.0%; Score 1365; DB 2; Length 2408;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGAAACGGGGGACTACAGAGACATGTTTGGCCATATTACAGCGCGTCACTG 60
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 457 ATGAAACGGGGGACTACAGAGACATGTTTGGCCATATTACAGCGCGTCACTG 516
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 61 CTAAAAATTCGAGACACAACAAAGCTCCTCAATTTAAAGTCCCAATTAATGATCA 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 517 CTAAAAATTCGAGACACAACAAAGCTCCTCAATTTAAAGTCCCAATTAATGATCA 576
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 GCAATCAAAAACATTTGATGGCAAAAAGGTATGATTAAGTACAGCACTTAATGATTA 180
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 577 GCAATCAAAAACATTTGATGGCAAAAAGGTATGATTAAGTACAGCACTTAATGATTA 636
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 181 GATGATGGATGATGGCCACTGCAAAAGCGTGTGCTAGTGGGCAAAATTAATGATTA 240
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 637 GATGATGGATGATGGCCACTGCAAAAGCGTGTGCTAGTGGGCAAAATTAATGATTA 696
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 241 TATGACATCGTCTCCGCTTTAGCAGGTGACCCCAAAAACAGTGTGATCTCCACTTCAT 300
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 697 TATGACATCGTCTCCGCTTTAGCAGGTGACCCCAAAAACAGTGTGATCTCCACTTCAT 756
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 301 TTATCTATCAAAAAGTGGTGTGATGATGATGACAGCTGGAATAATGCTGGAAGACTA 360
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 757 TTATCTATCAAAAAGTGGTGTGATGATGATGACAGCTGGAATAATGCTGGAAGACTA 816
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 361 TTTGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 817 TTTGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 421 TGGTACAGTCTGCTACTTTTAACCAAGATGGCCAGTCCGTTTATTTCTATACAGATTAC 480
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 877 TGGTACAGTCTGCTACTTTTAACCAAGATGGCCAGTCCGTTTATTTCTATACAGATTAC 936
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 481 TCAGTATCTCTGAAGATGGTGGAAACCGGTGCTGTATCAAAATCATTTCAACTGCTCAA 540
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 937 TCAGTATCTCTGAAGATGGTGGAAACCGGTGCTGTATCAAAATCATTTCAACTGCTCAA 996
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 541 GTAAACTTATCCAGCGGATGACGATCACTTAAGAGCGATGATGATGATGATGATGATGAT 600
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 997 GTAAACTTATCCAGCGGATGACGATCACTTAAGAGCGATGATGATGATGATGATGATGAT 1056
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 601 TCTGTCTTTGATGGCGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1057 TCTGTCTTTGATGGCGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 661 GGCAAGTGGATTCAGGTGATTAACCATCTTTAAGAGCGCTCACTATGTTGAAGATTAAG 720
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1117 GGCAAGTGGATTCAGGTGATTAACCATCTTTAAGAGCGCTCACTATGTTGAAGATTAAG 1176
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 721 GGCCATTAATATCTTGTCTTTGAAGCGAATCTGGAACACAGATGATGATCAAGCGCAT 780
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1177 GGCCATTAATATCTTGTCTTTGAAGCGAATCTGGAACACAGATGATGATCAAGCGCAT 1236
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 781 CAGTCTTCAATTAATAAGCTTACTATGCGGAAGTGAAGCTCTCTTCCAGAAATGAAAA 840
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1237 CAGTCTTCAATTAATAAGCTTACTATGCGGAAGTGAAGCTCTCTTCCAGAAATGAAAA 1296
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 841 AATAAATCGCTTCAAAAGCTTAAACCAATGCTTTTGAAGATGATGATGATGATGATGAT 900
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1297 AATAAATCGCTTCAAAAGCTTAAACCAATGCTTTTGAAGATGATGATGATGATGATGAT 1356
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 901 ATTGTGAATGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1357 ATTGTGAATGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 961 AACACAGTAGCAGATGAAGTGAAGCGGCCAATATATTAAATCAATTAATTAATGATGAT 1020
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1417 AACACAGTAGCAGATGAAGTGAAGCGGCCAATATATTAAATGAATTAATTAATGATGAT 1476
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1021 CTATTCACGATTCAGAGGATCCAAAATGACAGTGAATTAACGACAAAGATGTT 1080
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
DB 1477 CTATTCACGATTCAGAGGATCCAAAATGACAGTGAATTAACGACAAAGATGTT 1536
OY 1081 TATATGCTAGGCGCGGAGGAGGACTCCTTAATGAGCCACACACCGGATTAATGAAGT 1140
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1537 TATATGCTAGGCGCGGAGGAGGACTCCTTAATGAGCCACACACCGGATTAATGAAGT 1596
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1141 GGACTTGTATGAACATGAATCTTGACCCGCTGATCTCAACACACTTACTCATTTGC 1200
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1597 GGACTTGTATGAACATGAATCTTGACCCGCTGATCTCAACACACTTACTCATTTGC 1656
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1201 GGTATCCCGCACCCCTGAAGTAATATGTTGTTACTCAACAGTATATGAGCAATGAGGC 1260
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1657 GGTATCCCGCACCCCTGAAGTAATATGTTGTTACTCAACAGTATATGAGCAATGAGGC 1716
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1261 TTCTATCCAGAACATCACTCTCACCTGCGGAGCAAGCTTGAGTTAATTAAGGCT 1320
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1717 TTCTATCCAGAACATCACTCTCACCTGCGGAGCAAGCTTGAGTTAATTAAGGCT 1776
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1321 GACACATCTGGAGAGAAAATAGTTCCGAGACAGACATTTCCCA 1365
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1777 GACACATCTGGAGAGAAAATAGTTCCGAGACAGACATTTCCCA 1821
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
US-09-317-179-5
: Sequence 5, Application US/09317179
: Patent No. 6383769
GENERAL INFORMATION:
APPLICANT: Tsusaki et al.
TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE ACTIVITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: BROMDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/317, 179
FILING DATE: 24-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/870, 827
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BROMDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2408 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus sp.
INDIVIDUAL ISOLATE: V230 (FERM BP-5054)
FEATURE:
NAME/KEY: signal peptide
LOCATION: 1..360
IDENTIFICATION METHOD: E
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
```

US-09-317-179-5

Query Match 100.0%; Score 1365; DB 4; Length 2408;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAACAGGGGAGCTACAGAGAGCTATGTTTTGGCCCATATTTACACGGCCTGACATG 60
DB 457 ATGAACAGCGGGAGCTACAGAGAGAGCTATGTTTTGGCCCATATTTACACGGCCTGACATG 516
QY 61 CTAATAATTCAGAGACACAAACAGTCTCTCAATTTAAAGTCCCATTCATTCATGATCA 120
DB 517 CTAATAATTCAGAGACACAAACAGTCTCTCAATTTAAAGTCCCATTCATTCATGATCA 576
QY 121 GCAATCAAAAACATTTGATTCGCAAAAAGGATGATAGTACAGGACCTTAATAGATTGA 180
DB 577 GCAATCAAAAACATTTGATTCGCAAAAAGGATGATAGTACAGGACCTTAATAGATTGA 636
QY 181 GATGATGGGATAGCTGGCCACTGCAAAAACGCTGATGCTGCGCAAAATTTATCATGSA 240
DB 637 GATGATGGGATAGCTGGCCACTGCAAAAACGCTGATGCTGCGCAAAATTTATCATGSA 696
QY 241 TATCATGCTGCTCCGCTTTAGCAGGTACCCCAAAAACAGATGATACCTCATTCAT 300
DB 697 TATCATGCTGCTCCGCTTTAGCAGGTACCCCAAAAACAGATGATACCTCATTCAT 756
QY 301 TTATCTATCAAAAAGTGGTGTATCATGATTCAGCTGGAAGAAATGCTGGAAGAGTA 360
DB 757 TTATCTATCAAAAAGTGGTGTATCATGATTCAGCTGGAAGAAATGCTGGAAGAGTA 816
QY 361 TTTGAAGATATGATTAATTTGTTCCAAATGATCCGATCTTTAAATTCACACAGAG 420
DB 817 TTTGAAGATATGATTAATTTGTTCCAAATGATCCGATCTTTAAATTCACACAGAG 876
QY 421 TGGTAGGTTCTGCTACTTTAAACCAAGATGCCAAGTCCGTTTCTATACAGATTAC 480
DB 877 TGGTAGGTTCTGCTACTTTAAACCAAGATGCCAAGTCCGTTTCTATACAGATTAC 936
QY 481 TCAGTAAATCTGTAAGATGTGGAACCGTCTGTAACCAAAATCATTTCAACTGCTAA 540
DB 937 TCAGTAAATCTGTAAGATGTGGAACCGTCTGTAACCAAAATCATTTCAACTGCTAA 996
QY 541 GTAACTATATCCACCGCGATGACCTAATTAAGTCGATGAGTATGTCATATAA 600
DB 997 GTAACTATATCCACCGCGATGACCTAATTAAGTCGATGAGTATGTCATATAA 1056
QY 601 TCTGCTTTGATGGGAGAGGATACAGTTATCAAAATTTACAGCAATTTATCGATGAA 660
DB 1057 TCTGCTTTGATGGGAGAGGATACAGTTATCAAAATTTACAGCAATTTATCGATGAA 1116
QY 661 GGCAAGTGAATTTAGTGTATACCATCTTTAAGAGACCTCACTATATGGAAGATAG 720
DB 1117 GGCAAGTGAATTTAGTGTATACCATCTTTAAGAGACCTCACTATATGGAAGATAG 1176
QY 721 GGCCATTAATATCTTGTCTTTGAAAGCAATCTGGAACAACAGATGTTATCAAGGCAT 780
DB 1177 GGCCATTAATATCTTGTCTTTGAAAGCAATCTGGAACAACAGATGTTATCAAGGCAT 1236
QY 781 CAGTCTTCATTAATTAAGTCTTACTATGCGGAAGTGAAGTCTCTTCCAGAAATGAAA 840
DB 1237 CAGTCTTCATTAATTAAGTCTTACTATGCGGAAGTGAAGTCTCTTCCAGAAATGAAA 1296
QY 841 AATAAACTGCTTCAAAAGTCTTAATAAACAAATGCTTTAGGAAAGTGTATGAGCATCA 900
DB 1297 AATAAACTGCTTCAAAAGTCTTAATAAACAAATGCTTTAGGAAAGTGTATGAGCATCA 1356
QY 901 ATTGTTGAATTTGGCGATGACTATACAGTGAAGTGTATGAAACCATTAAGTCATCA 960
DB 1357 ATTGTTGAATTTGGCGATGACTATACAGTGAAGTGTATGAAACCATTAAGTCATCA 1416
QY 961 AACACAGTACAGATGATGAGTGAAGCGCCCAATATATTAAATGAATTAATGATAT 1020
DB 1417 AACACAGTACAGATGATGAGTGAAGCGCCCAATATATTAAATGAATTAATGATAT 1476

```

```

QY 1021 CTATTACAGGATTTCAAGAGATCCAAATGACAGATGATGAAATTAACGACAAAGATGTT 1080
DB 1477 CTATTACAGGATTTCAAGAGATCCAAATGACAGATGATGAAATTAACGACAAAGATGTT 1536
QY 1081 TATATGCTAAGGCGCCGAGAGGAGTCTTAAATGAGCCACAGAACCCGATTAATGAAACT 1140
DB 1537 TATATGCTAAGGCGCCGAGAGGAGTCTTAAATGAGCCACAGAACCCGATTAATGAAACT 1596
QY 1141 GGACTTGTATTAATGATATCTTATGACCTGCTGATCTCACACACACTTACTTCATTGC 1200
DB 1597 GGACTTGTATTAATGATATCTTATGACCTGCTGATCTCACACACACTTACTTCATTGC 1656
QY 1201 GGTATCCCGCACCCGAGAGTAAATATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1657 GGTATCCCGCACCCGAGAGTAAATATGATGATGATGATGATGATGATGATGATGATGAT 1716
QY 1261 TTTATCCAGAACATCACTCTCACCTGCGGAGCAAGCTTGGGTTAATATTAAGGCTCT 1320
DB 1717 TTTATCCAGAACATCACTCTCACCTGCGGAGCAAGCTTGGGTTAATATTAAGGCTCT 1776
QY 1321 GACACATCTGGAGGAGAAATAGTTCCGACACAGACATTCCCA 1365
DB 1777 GACACATCTGGAGGAGAAATAGTTCCGACACAGACATTCCCA 1821

```

RESULT 5

US-09-058-746-1/c

Sequence 1, Application US/09058746

Patent No. 6022716

GENERAL INFORMATION:

APPLICANT: Ilya Chumakov

APPLICANT: Hiroaki Tanaka

TITLE OF INVENTION: High Throughput DNA sequencing vector

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESS: Knobb, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/058, 746

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: GENSET.015A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 10317 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: DOUBLE

TOPOLOGY: CIRCULAR

MOLECULE TYPE: synthetic DNA

ORIGINAL SOURCE:

ORGANISM: Cloning vector pGendel

FEATURE:

NAME/KEY: pGendel

LOCATION: 1..10317

FEATURE:

NAME/KEY: Homology with X06404 comp1 (411..1668)

LOCATION: 9..1266

IDENTIFICATION METHOD: blastn against X06404
FEATURE:
NAME/KEY: Kanamycin resistance gene CDS
LOCATION: 142..957
IDENTIFICATION METHOD: By homology to X06404
FEATURE:
NAME/KEY: Tn1000's right end
LOCATION: complement 1332..1371
IDENTIFICATION METHOD: blastn against X06200
FEATURE:
NAME/KEY: Homology with U46017 (1-472)
LOCATION: 1423..1894
IDENTIFICATION METHOD: blastn against U46017
FEATURE:
NAME/KEY: single stranded DNA replication origin
LOCATION: 1423..1894
IDENTIFICATION METHOD: By homology to U46017
OTHER INFORMATION: mutation T -> C 1658
FEATURE:
NAME/KEY: Homology with U51113 (2382..6997)
LOCATION: 1896..6544
IDENTIFICATION METHOD: blastn against U51113
FEATURE:
NAME/KEY: OriS
LOCATION: 1972..2188
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: repELR
LOCATION: 2897..2918
OTHER INFORMATION: Described in seqid 16
FEATURE:
NAME/KEY: Repe
LOCATION: 2903..3034
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: T3
LOCATION: 3043..3059
OTHER INFORMATION: Described in seqid 17
FEATURE:
NAME/KEY: LRT3RA
LOCATION: complement 3045..3069
OTHER INFORMATION: Described in seqid 15
FEATURE:
NAME/KEY: Incc
LOCATION: 3070..3320
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: insertion 33 bases 3038..3071
FEATURE:
NAME/KEY: Para
LOCATION: 3655..4821
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: mutation G -> A 3878
FEATURE:
NAME/KEY: ParaB
LOCATION: 4821..5792
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: ParaC
LOCATION: 5865..6382
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: Homology with J01688 (complement 175..819)
LOCATION: 6574..7218
IDENTIFICATION METHOD: blastn against J01688
OTHER INFORMATION: mutation A -> G 7096
FEATURE:
NAME/KEY: CDS streptomycin sensitivity gene
LOCATION: complement 6716..7090
IDENTIFICATION METHOD: By homology to J01688
OTHER INFORMATION: mutation A -> G 6728
OTHER INFORMATION: mutation G -> C 6821
OTHER INFORMATION: mutation C -> T 6866
OTHER INFORMATION: mutation T -> C 7013

OTHER INFORMATION: mutation T -> A 7058
FEATURE:
NAME/KEY: rpsLR
LOCATION: 7155..7174
OTHER INFORMATION: Described in seqid 12
FEATURE:
NAME/KEY: SP6
LOCATION: 7230..7248
OTHER INFORMATION: Described in seqid 13
FEATURE:
NAME/KEY: Tn1000's left end
LOCATION: 7252..7291
IDENTIFICATION METHOD: blast (X06200)
FEATURE:
NAME/KEY: Homology with X02730 (complement 37..1959)
LOCATION: 7305..9227
IDENTIFICATION METHOD: blastn against X02730
FEATURE:
NAME/KEY: CDS levansucrase gene
LOCATION: complement 7379..8800
IDENTIFICATION METHOD: By homology to X02730
OTHER INFORMATION: mutation T -> C 7466
OTHER INFORMATION: mutation A -> G 7739
OTHER INFORMATION: mutation T -> C (Asn -> Asp) 8347
OTHER INFORMATION: mutation T -> C 8600
OTHER INFORMATION: mutation G -> A (Ala -> Val) 8772
FEATURE:
NAME/KEY: SLR3
LOCATION: 8711..8731
OTHER INFORMATION: Described in seqid 14
FEATURE:
NAME/KEY: Homology with J01636 (complement 1158..1465)
LOCATION: 9298..9623
IDENTIFICATION METHOD: blastn against J01636
FEATURE:
NAME/KEY: CDS alpha peptide beta-galactosidase
LOCATION: complement 9276..9497
IDENTIFICATION METHOD: By homology to J01636
FEATURE:
NAME/KEY: primer HE1
LOCATION: complement 9465..9479
FEATURE:
NAME/KEY: primer HE2
LOCATION: 9461..9475
FEATURE:
NAME/KEY: primer lacZ52AVT
LOCATION: complement 9603..9630
FEATURE:
NAME/KEY: primer lacZ2Mlu
LOCATION: 9289..9314
FEATURE:
NAME/KEY: Homology with M77789 (1889..2576)
LOCATION: 9629..10315
IDENTIFICATION METHOD: blastn against M77789
FEATURE:
NAME/KEY: high copy-number double-stranded DNA replication origin
LOCATION: complement 9629..10315
IDENTIFICATION METHOD: By homology to M77789
OTHER INFORMATION: mutation C -> T 9803
OTHER INFORMATION: site ScaI 10029 - 10034
OTHER INFORMATION: site PmlI 10038 - 10043
OTHER INFORMATION: CLONING SITES 10031 -- 10041
FEATURE:
NAME/KEY: oriLRd
LOCATION: 9856..9881
OTHER INFORMATION: Described in seqid 8
FEATURE:
NAME/KEY: OSI
LOCATION: 10009..10026
OTHER INFORMATION: Described in seqid 10
FEATURE:
NAME/KEY: ORI
LOCATION: complement 10046..10062

OTHER INFORMATION: Described in seqid 11
 FEATURE:
 NAME/KEY: ORILRT
 LOCATION: complement 10182..10202
 OTHER INFORMATION: Described in seqid 9
 US-09-058-746-1

Query Match 42.3%; Score 577; DB 3; Length 10317;
 Best Local Similarity 67.2%; Pred. No. 3.5e-152;
 Matches 902; Conservative 0; Mismatches 390; Indels 51; Gaps 4;

QY 1 ATGACAGCGGGGACTACAAAGAGACTATGTTTCCCATATTACACGCGCTGACATG 60
 DB 8707 ACGAACCAAAAGCCATATAGGAAGAAACATAGCGCATTTCCATATATACACCGCATGATG 8648
 QY 61 CTAAAAATTCAGAGACAAACAAACAGTCTCAATTTAAAGGCGCTCATTCATCATGATCA 120
 DB 8647 CTGCAATTCCTGACAGCAGCAAAAAATGAATAATATCAAGTTCTTGAGTTGATTCGCTC 8588
 QY 121 GCAATCAAAACATTTGATGCGCAAAAGGTATGATAGTCAAGCACTTAATGATTTA 180
 DB 8587 ACAATTAATAATCTCTCTCGCAAAAG-----CCTG 8555
 QY 181 GATGATAGGATAGCTGCGCACATGCAAAAGCTGATGCTACTGCGCAATTAATCATGGA 240
 DB 8554 GAGCTTTGGAGACGCTGCGCATTAACAAACGCTGACGCACTGTGCAAACTATACAGGC 8495
 QY 241 TATCACATGCTGCCCTTATAGCAGTACGCCCAAAAAACAGATGATATCCACTCAT 300
 DB 8494 TACCACATCTCTTTGATTAACCGGAGATCTTAATAATGGGATGACACATCATTTAC 8435
 QY 301 TTATTTATCAAAAAGTGGTGTATACATGATGACAGCTGGAATAATGCTGGAAGATA 360
 DB 8434 ATGTTTATCAAAAAGTGGGCAAACTTATGACAGCTGGAATAATGCTGCGCGGCTC 8375
 QY 361 TTGGAATATGATTAATTTGTTCCAAATGATCCGATCTTAATATCAACACAGAG 420
 DB 8374 TTTTAAAGACGCAAAATTCAGATGAGATTCCTTAAGAACCAACCAAGAA 8315
 QY 421 TGGCAGGTTCTGCTACTTTAACCAAGATGGCCAAAGCTTATCTTATACAGATTAC 480
 DB 8314 TGGCAGGTTCTGACACATTTACATCTGACGGAATAATCCGTTTATCTACACTGATTTC 8255
 QY 481 TCAGATTAATCTGAAGATGTGGAACCGGTGCTGTGAACCAAAATCATTTCAACTGTCAA 540
 DB 8254 TCCGGTAAACATTA-----CGGCAACAAACACATGACACAGTCAAA 8213
 QY 541 GTAAACTTATCCAGCGCGGATGACACTTAAAGTCGATGAGATATGATCATAAA 600
 DB 8212 GTTACGATATCAGCATGACAGCTCT--TTGAACATCAACGGTGTAAAGGATTATAAA 8156
 QY 601 TCTGCTTTGATGGCGGAGACGCTACAGTTTATCAAAATATTCAGCAATTTATGATGAA 660
 DB 8155 TCAATCTTTA---CGGTGACGGAACAAACGATATCAAAATGTACAGCGATTATGATGAA 8099
 QY 661 GCGAAGTGAATTCAGGTGATTAACCATATCTTTAGAGACCCCTCACTATGTTGAAGATAG 720
 DB 8098 GCGAAGTGAATTCAGGTGATTAACCATATCTTTAGAGACCCCTCACTATGTTGAAGATAG 8039
 QY 721 GGGCATTAATATCTGCTTTGGAAGCAATACAGTGAACACAGATGTTTCAAGGCGAT 780
 DB 8038 GGGCACAATAATCTGATTTTGAAGCAACAGTGAACCTGAAGTGTCTACCAAGGGGAA 7979
 QY 781 CAGTCTTTCAATATAAGCTTACTATGCGGAAGTACGCTCTTCGCAAGATGAAA 840
 DB 7978 GATCTTTATTTAAACAAAGCATATGCGCAAAAGCATATCTCTTCGCAAGAAAGT 7919
 QY 841 AATTAAGCTTCAAAAGCTCTAAACAAATTCCTTTAGGAAGTGTGATGAGC 900
 DB 7918 CAAAAGCTTCTGCAAGCGATTAACAAACGACGCTGATAGCAACGGGCGCTCTGCT 7859
 QY 901 ATTGTTAATTTGGCCGATGACGTAGTGAAGAAAGTTTATGAACCATTAAGTGCATCA 960

DB 7858 ATGATGAGCTAAACGATGATTACACACGTGAAAAAGTGAATAACCGCTGATTGATCT 7799
 QY 961 AACACAGTACGATGATGAGTGCAGACGGCCCAATATATTAAATGANTAAATGAT 1020
 DB 7798 AACACAGTACGATGATGAGTGCAGACGGCCCAATATATTAAATGANTAAATGAT 7739
 QY 1021 CTATTCACGATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 7738 CTGTTCACTGATCTCCCGGATCAAAATGATGATGATGATGATGATGATGATGAT 7679
 QY 1081 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 7678 TATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7619
 QY 1141 GCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 7618 GCGCTTGTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7559
 QY 1201 GGTATCCGCAACCTGGAAGTAAATGATGATGATGATGATGATGATGATGATGAT 1260
 DB 7558 GCTGTACCTCAAGGAAGAAACATGATGATGATGATGATGATGATGATGATGAT 7499
 QY 1261 TTTATCCAGACATCTCTACCTGCGGACAGCTTGAGGTTAATTAAGGCTCT 1320
 DB 7498 TTTATCCAGACATCTCTACCTGCGGACAGCTTGAGGTTAATTAAGGCTCT 1320
 QY 1321 GACATCTGGAAGGAATAATAG 1343
 DB 7438 AAAACATCTGTTTCAAGACAG 7416

RESULT 6

US-09-438-142-1/c
 Sequence 1, Application US/09438142

Patent No. 6258571

GENERAL INFORMATION:

APPLICANT: Ilya Chumakov

APPLICANT: Hiroaki Tanaka

TITLE OF INVENTION: High Throughput DNA sequencing vector

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knodbe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/438,142

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Israelson, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: GENSET.015C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-0176

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 10317 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: synthetic DNA

ORGANISM: Cloning vector pGendEL

FEATURE: NAME/KEY: pGendel

LOCATION: 1..10317
 FEATURE:
 NAME/KEY: Homology with X06404 compl (411..1668)
 LOCATION: 9..1266
 IDENTIFICATION METHOD: blastn against X06404
 FEATURE:
 NAME/KEY: Kanamycin resistance gene CDS
 LOCATION: 142..957
 IDENTIFICATION METHOD: By homology to X06404
 FEATURE:
 NAME/KEY: Tn1000's right end
 LOCATION: complement 1332..1371
 IDENTIFICATION METHOD: blastn against X60200)
 FEATURE:
 NAME/KEY: Homology with U46017 (1-472)
 LOCATION: 1423..1894
 IDENTIFICATION METHOD: blastn against U46017
 FEATURE:
 NAME/KEY: single stranded DNA replication origin
 LOCATION: 1423..1894
 IDENTIFICATION METHOD: By homology to U46017
 OTHER INFORMATION: mutation T -> C 1658
 FEATURE:
 NAME/KEY: Homology with U51113 (2382..6997)
 LOCATION: 1896..6544
 IDENTIFICATION METHOD: blastn against U51113
 FEATURE:
 NAME/KEY: Orls
 LOCATION: 1972..2188
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: repELR
 LOCATION: 2897..2918
 OTHER INFORMATION: Described in seqid 16
 FEATURE:
 NAME/KEY: RepE
 LOCATION: 2903..3034
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: T3
 LOCATION: 3043..3059
 OTHER INFORMATION: Described in seqid 17
 FEATURE:
 NAME/KEY: LRT3RA
 LOCATION: complement 3045..3069
 OTHER INFORMATION: Described in seqid 15
 FEATURE:
 NAME/KEY: Incc
 LOCATION: 3070..3320
 IDENTIFICATION METHOD: By homology to U51113
 OTHER INFORMATION: Insertion 33 bases 3038..3071
 FEATURE:
 NAME/KEY: Para
 LOCATION: 3655..4821
 IDENTIFICATION METHOD: By homology to U51113
 OTHER INFORMATION: mutation G -> A 3878
 FEATURE:
 NAME/KEY: Parab
 LOCATION: 4821..5792
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: ParC
 LOCATION: 5865..6382
 IDENTIFICATION METHOD: By homology to U51113
 NAME/KEY: Homology with J01688 (complement 175..819)
 LOCATION: 6574..7218
 IDENTIFICATION METHOD: blastn against J01688
 OTHER INFORMATION: mutation A -> G 7096
 FEATURE:
 NAME/KEY: CDS streptomycin sensitivity gene
 LOCATION: complement 6716..7090
 IDENTIFICATION METHOD: By homology to J01688

OTHER INFORMATION: mutation A -> G 6728
 OTHER INFORMATION: mutation G -> C 6821
 OTHER INFORMATION: mutation C -> T 6866
 OTHER INFORMATION: mutation T -> C 7013
 OTHER INFORMATION: mutation T -> A 7058
 FEATURE:
 NAME/KEY: rpsLR
 LOCATION: 7155..7174
 OTHER INFORMATION: Described in seqid 12
 FEATURE:
 NAME/KEY: SP6
 LOCATION: 7230..7248
 OTHER INFORMATION: Described in seqid 13
 FEATURE:
 NAME/KEY: Tn1000's left end
 LOCATION: 7252..7291
 IDENTIFICATION METHOD: blast (X60200)
 FEATURE:
 NAME/KEY: Homology with X02730 (complement 37..1959)
 LOCATION: 7305..9227
 IDENTIFICATION METHOD: blastn against X02730
 FEATURE:
 NAME/KEY: CDS levansucrase gene
 LOCATION: complement 7379..8800
 IDENTIFICATION METHOD: By homology to X02730
 OTHER INFORMATION: mutation T -> C 7466
 OTHER INFORMATION: mutation A -> G 7739
 OTHER INFORMATION: mutation T -> C (Asn -> Asp) 8347
 OTHER INFORMATION: mutation T -> C 8600
 OTHER INFORMATION: mutation G -> A (Ala -> Val) 8772
 FEATURE:
 NAME/KEY: SLR3
 LOCATION: 8711..8731
 OTHER INFORMATION: Described in seqid 14
 FEATURE:
 NAME/KEY: Homology with J01636 (complement 1158..1465)
 LOCATION: 9298..9623
 IDENTIFICATION METHOD: blastn against J01636
 FEATURE:
 NAME/KEY: CDS alpha peptide beta-galactosidase
 LOCATION: complement 9276..9497
 IDENTIFICATION METHOD: By homology to J01636
 FEATURE:
 NAME/KEY: Primer HE1
 LOCATION: complement 9465..9479
 FEATURE:
 NAME/KEY: Primer HE2
 LOCATION: 9461..9475
 FEATURE:
 NAME/KEY: primer LacRS2AVr
 LOCATION: complement 9603..9630
 FEATURE:
 NAME/KEY: primer LacE2Mlu
 LOCATION: 9289..9314
 FEATURE:
 NAME/KEY: Homology with M77789 (1889..2576)
 LOCATION: 9629..10315
 IDENTIFICATION METHOD: blastn against M77789
 FEATURE:
 NAME/KEY: high copy-number double-stranded DNA replication origin
 LOCATION: complement 9629..10315
 IDENTIFICATION METHOD: By homology to M77789
 OTHER INFORMATION: mutation C -> T 9803
 OTHER INFORMATION: site ScaI 10029 - 10034
 OTHER INFORMATION: site PmlI 10038 - 10043
 OTHER INFORMATION: CLONING SITES 10031 - 10041
 FEATURE:
 NAME/KEY: oriLRd
 LOCATION: 9856..9881
 OTHER INFORMATION: Described in seqid 8
 FEATURE:
 NAME/KEY: OS1
 LOCATION: 10009..10026

```

OTHER INFORMATION: Described in segid 10
FEATURE:
NAME/KEY: OR1
LOCATION: complement 10046..10062
OTHER INFORMATION: Described in segid 11
FEATURE:
NAME/KEY: orlRr
LOCATION: complement 10182..10202
OTHER INFORMATION: Described in segid 9
US-09-438-142-1

Query Match      42.38; Score 577; DB 4; Length 10317;
Best Local Similarity 67.28; Pred. No. 3,5e-152;
Matches 902; Conservative 0; Mismatches 390; Indels 51; Gaps 4;

OY 1 ATGAAAGCGGGGACCTACAGGAAGACTATGTTGGCCATATTACACGGCGACATG 60
DB 8707 ACGAACCAAAAGCCATATAGGAACATACGGCATTTCCATATTACCGCCATGATATG 8648
OY 61 CTAAAAATTCAGAGACAACAACAGTCCTCAATTTAAAGTGCCTCAATTCATCA 120
DB 8647 CTGCAATTCCTGAAAGCAAAAAAATGAAATATCAAGTCTCGAGTTGATTCGTC 8588
OY 121 GCAATCAAAACATTGATTGGCAAAAGGATATGATAGTCAGCAACTAATGATTTA 180
DB 8587 ACAATTTAAATATCTCTTCGCAAAAG-----CCTG 8555
OY 181 GATGTATGGATCTACCTGCGCCACGCAAAAGCTGATGCTAGTGGCAATATCATGGA 240
DB 8554 GACCTTTGGGACACTGGCCCTTACAAACGCTACGCGCATCTGCGCAACTATACGGC 8495
OY 241 TATCACTGCTGCTCCGCTTACAGAGTGACCAAAAAACAAGTATGATCTCACTTCAT 300
DB 8494 TACCACATCGCTTTCGATTAAGCGGAGATCTTAAATAGCGGATGACATTCATTTAC 8435
OY 301 TTATCTATCAAAAAGTCGCTGATACATCATTTACAGCTGGAAAAATGCTGGAAGCTA 360
DB 8434 ATGTTCTATCAAAAAGTCGCGGAACTTCTATTACAGCTGGAAAAAGCTGCGCGCTC 8375
OY 361 TTGAAGATATGATTAATTTGTCCAAATGATCCGATCTTAATATACAAACAGAGAG 420
DB 8374 TTTTAAACACGCGCAAAATTCATGACAGATGATTCCTTAAAGACCAACCAACAGAA 8315
OY 421 TGGTCAGGTTCTGCTACTTTAAACCAAGATGGCCAGTCCGTTATTCATACAGATTAC 480
DB 8314 TGGTCAGGTTCCAGCACATTTTACATCTGAGGAAAAATCCGTTATTCTACAGATTTC 8255
OY 481 TCAGGTAATCTGGAAGATGGTGGAAACCGGTGCTGTAACCAATCATTTCACTGCTCAA 540
DB 8254 TCCGCTAAACATTA-----CGGCAAAACAACACTGACAACTGCACAA 8213
OY 541 GTAAACTTATCCGCGCGGATGACGTAACCTTAAGTCGATGATTCGATCATGAA 600
DB 8212 GTTACGTAATCAGCATAGACAGCTCT--TTGACATCAACGGGTGAGAGATTAATAA 8156
OY 601 TCTGCTTTGATGGCGAGAGCGGTACAGATTATCAAAATATTCGCAATTTATGATGAA 660
DB 8155 TCATCTTTGA---CGGTGAAGGAAAAACGTAATCAAAATGTACAGAGTTATGATGAA 8099
OY 661 GCGAAGTGGATTGAGTGAATACCATCTTTAAAGACCTCTACTATGTTGAAGATPAG 720
DB 8098 GCGAATCATACGCTCAGCGACAAACCATACGCTGAGAGATCTCTACTAGAGATPAA 8039
OY 721 GGCATTAATATCTGCTTTGAGGGAATACGTGAACAACAGATGGTTATCAGGGGAT 780
DB 8038 GCGCACAATACTAGTATTGAAAGCAAAACCTGGAATGAGATGGCTACCAAGCGCAA 7919
OY 781 CAGTCTTTCAATAATAAGCTTACTATGGGGAAGTACGCTCTTCTCCAGAAATGAAAA 840
DB 7918 GAATCTTTATTAACAAGACTACTATGCAAAAGCATCATCTTCCGTCAGAAAGT 7919
OY 841 AATTAATGCTTCAAAAGTCTTAAAAACAATTCCTTTAGGAAAGGTGCAATTTAGGC 900
DB 841 AATTAATGCTTCAAAAGTCTTAAAAACAATTCCTTTAGGAAAGGTGCAATTTAGGC 900

```

```

DB 7918 CAAAACTTTCGCAAGCGATTAATAAAACGACGGCTGAGTTAGCAAAAGCGGCTCTCGGT 7859
OY 901 AATGTTGAATGGCGGATGATATACAGTGAAGAGTGTATGAACCAATTAGTGCATCA 960
DB 7858 ATGATTGAGCTTAACAGATGATATTACACTGTAAGAAAGTATGAACCGCTATTGCACT 7799
OY 961 AACACAGTACAGATGAAGTGAAGCGCCCAATATATTAAATGAATGAATGAATGATAT 1020
DB 7798 AACACAGTACAGATGAATGAATGAAGCGCCCAATCTTTAAATGAACCGCAATGATAT 7739
OY 1021 CTATTACAGGATTCAGAGATTCACAAATGACAGTGAATGAATGAATGAATGAATGAT 1080
DB 7738 CTGTTCACTGATCTCCGCGGATCAAAATGACAGATGAGTGAAGCATTAAGTATGAT 7679
OY 1081 TATATGCTAGGCGCCGCGGACCTCTTAATATGAGCCCAACCAACCGATTAATGAAC 1140
DB 7678 TACATGCTGTTATGTTTCTTAATTTCTTAACCTGACCATACAGCGCTGACAAACT 7619
OY 1141 GGACTTGTATTAACATGAATCTTGACCTGCTGATCTGACACACACTTACTCTCATTTG 1200
DB 7618 GGCCTGTGTTAAATGATCTTGATCTTAACGATGTAACCTTACTTACTTACTTACTT 7559
OY 1201 GGTATCCGCGACCTGGAAGTAAATGATGTTACTACAGATTAATGACGAATGAGGC 1260
DB 7558 GCTGTACTCTCAAGCGGAAAGAAATGTCGTGATTAACACTATATGACAAACAGAGA 7499
OY 1261 TTCTATCGAAGAACATCTCTCACTGCGGAGCAACAGCTGGGTTAATTAAGGCTCT 1320
DB 7498 TTCTACGAGCAACAAATCAACGTTTGGCGGACCTTCTGCTGAACTAAGGCAAG 7439
OY 1321 GACACATCTGGAGAGAAATAG 1343
DB 7438 AAACATCTGTTGCAAAAGACAG 7416

```

```

RESULT 7
US-09-068-043-1
; Sequence 1, Application US/09068043.
; Patent No. 6048694
;
GENERAL INFORMATION:
APPLICANT: MICHAEL GENE BRAMUCCI
APPLICANT: VASANTHA NAGARAJAN
TITLE OF INVENTION: POSITIVE SELECTION
TITLE OF INVENTION: VECTOR FOR BACILLUS SP.
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND
ADDRESS: COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT OFFICE 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068.043
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,201
FILING DATE: NOVEMBER 3, 1995
ATTORNEY/AGENT INFORMATION:
NAME: FELTHAM, S. NEIL
REGISTRATION NUMBER: 36,506
REFERENCE/DOCKET NUMBER: CR-9807
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-6460
TELEFAX: 302-773-0164
;
; INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:
 LENGTH: 3305 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 US-09-068-043-1

Query Match 39.5%; Score 539.6; DB 3; Length 3305;
 Best Local Similarity 65.7%; Pred. No. 6,9e-142;
 Matches 873; Conservative 0; Mismatches 404; Indels 51; Gaps 4;

```

16 TACAAGAGAGATAGTGTGTTTCCCATATATACGCGCGAGACGTTAAATTCACAGA 75
1913 TACAAGAGAGATAGTGTGTTTCCCATATATACGCGCGAGACGTTAAATTCACAGA 1972
76 CACCAAAACAGTCTCATATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAACATT 135
1973 CAGCAGCAAAACAGTCTCATATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAACATT 2032
136 GATTGGCGCAAAAGGATGATAGTCAAGCACTTAATGATTTAGATGATGATGATGATG 195
2033 GAGTCTGCAAAAGGACTT-----GATGTGTCGACAGC 2065
196 TGGCCACTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 255
2066 TGGCCACTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2125
256 GCTTTAGCAGGTGACCCCAAAACAGTGTATCTCCATCTTCAATTTATTCATCAAAA 315
2126 GCTTTAGCAGGTGACCCCAAAACAGTGTATCTCCATCTTCAATTTATTCATCAAAA 2185
316 GTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
2186 GTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2245
376 AAATTTGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
2246 AAGTTTCGACGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2305
436 ACTTTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
2306 ACTTTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2364
496 GATGTGGAACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 555
2365 -----GCGCAAAACAGGCTGACACAGCCAGGTAAATGTGCACAA 2407
556 CCGGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
2408 TCTGAT---GACACACTCAAAATCAACGAGTGAAGATCACAAAACGATTTTGA--- 2461
616 GGAGAGCGTACAGTTATCAAAATATTCGCAATTTATGATGATGATGATGATGATGATG 675
2462 GGAGAGCGTACAGTTATCAAAATATTCGCAATTTATGATGATGATGATGATGATGATG 2521
676 GGTGATTAACCATCTTTAAGAGACCTCTATGTTGAAGATGAAGGACCATTAATATCT 735
2522 GCGGACACCATACGCTGAGAGACCTCTATGTTGAAGATGAAGGACCATTAATATCT 2581
736 GTCCTTGAAGCAATATCTGGAACAACAGATGTTATCAAGGCGATCAGTCTTTCATAT 795
2582 GTATTGCAAGCAACAGGAGCAAGAAACGATACCAAGGAGGAATCTTTATTTAAC 2641
796 AAGCTTACTATGCGGGAAGTACGCTCTTCTCCAGATGAAGAAAAATTAACGCTTCA 855
2642 AAGGCTTACTATGCGGGAAGTACGCTCTTCTCCAGATGAAGAAAAATTAACGCTTCA 2701
856 ACTCCATAAAACCAATTTGCTTTAGGATGATGATGATGATGATGATGATGATGATGATG 915
2702 AGCGCTAAAAACCGGATGCTGATTTAGGAAAGGCGCCCTGATGATGATGATGATGAT 2761

```

```

QY 916 GATGACTATACAGTGAAGAGTGTATGAACCATTAAGTCGATCAACACAGTAGCAGAT 975
DB 2762 AATGATTTACATTTAAAGAAAGTATGAAGCGGCTGATCAACACAGGTATAGAT 2821
QY 976 GAAGTCGACGCGCCCAATATATTTAAATGAATGAATATGATGATGATGATGATGATG 1035
DB 2822 GAATTCGACGCGCCCAATATATTTAAATGAATGAATGAATGATGATGATGATGATG 2881
QY 1036 AGAGATTCCAAAATGACGAGTATGATGATGATGATGATGATGATGATGATGATGATG 1095
DB 2882 CGCGTTTAAATGACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2941
QY 1096 GGAGGCGACTCTTAAATGCGCACACACACGATGAATGAAGTGAATGATGATGATGATGATG 1155
DB 2942 GATCAAAACCTTTTAAACCGCGCTTAAACGCGGCTGAACGCAACAGGCGTGTGCTGCA 3001
QY 1156 ATGAATCTTGAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
DB 3002 ATGGTCTTGAATCCAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 3061
QY 1216 GAAGTAAATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275
DB 3062 AAAGCAACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3121
QY 1276 CACTCTACCTGCGGACCAAGCTTGGGTTAATATTAAGGCTGACACATCTGAGAGA 1335
DB 3122 AAGGCAACATTTGGCGCAAGCTTCTTAATCAACATCAAAAGCAATAAACATCGTGTG 3181
QY 1336 GAAATAG 1343
DB 3182 AAAACAG 3189

```

RESULT 8

US-08-232-463-14/C
 ; Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHNEIDER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300,6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)836-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

[illegible]

```

FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7101 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7098
US-08-405-496A-9

Query Match 2.7%; Score 36.6; DB 2; Length 7101;
Best Local Similarity 60.6%; Pred. No. 1.5;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 166 AACTTAATGATTTAGATGATGATGCGATGCGCAAAAGCTGATGGTACTGCG 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3478 AATTCATAGTTTGTAGCTGAATGCGAATGCGAGAAATGGAAGGTGCTGACGTCACT 3537

QY 226 GCAATATATCATGATATACATGCTGCTCGGCTTTTACCA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3538 GTAACGTGATATATAGATCACTTTCTTTTACGACACATCA 3576

RESULT 12
US-08-915-136-9
Sequence 9, Application US/08915136
Patent No. 6290960
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESS: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907

```

FLING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FLING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FLING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7101 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7098
US-08-915-136-9

Query Match 2.7%; Score 36.6; DB 4; Length 7101;
Best Local Similarity 60.6%; Pred. No. 1.5;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 166 AACTTAATAGATTAGATGATGGAGTAGCTGCCACGCAAAACGCTGATGTAATCGG 225
DB 3478 AATTCATAGATTAGTGAATGGAATCTGGAAGATGGAAGGCGTTCAGGTCATCT 3537
QY 226 GCAATTAATCATGATATCATCATCTCTCCGCTTTAGCA 264
DB 3538 GTAACGTATGATATGATACCTCTTTTCAGCACCATCA 3576

RESULT 13
US-08-957-310-9
Sequence 9, Application US/08957310
Patent No. 6365158
GENERAL INFORMATION:
APPLICANT: Williams, James A.
APPLICANT: Kirk, John A.
TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
TITLE OF INVENTION: OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,310
FLING DATE: 23-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FLING DATE: 24-OCT-1994
APPLICATION NUMBER: US 08/161,907
FLING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321

FLING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FLING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7101 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7098
US-08-957-310-9

Query Match 2.7%; Score 36.6; DB 4; Length 7101;
Best Local Similarity 60.6%; Pred. No. 1.5;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 166 AACTTAATAGATTAGATGATGGAGTAGCTGCCACCTGCAAAACGCTGATGTAATCGG 225
DB 3478 AATTCATAGATTAGTGAATGGAATCTGGAAGATGGAAGGCGTTCAGGTCATCT 3537
QY 226 GCAATTAATCATGATATCATCATCTCTCCGCTTTAGCA 264
DB 3538 GTAACGTATGATATGATACCTCTTTTCAGCACCATCA 3576

RESULT 14
5231168-1
Patent No. 5231168
APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
TITLE OF INVENTION: MALARIA ANTIGEN
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/409,658
FLING DATE: 18-SEP-1989
SEQ ID NO: 1
LENGTH: 3095
5231168-1

Query Match 2.7%; Score 36.4; DB 6; Length 3095;
Best Local Similarity 42.0%; Pred. No. 1.2;
Matches 211; Conservative 0; Mismatches 291; Indels 0; Gaps 0;
QY 477 TTACTCAGTAACTCCGAGATGATGGAACCGGATGATGTAACCAATCTTCACTGC 536
DB 120 TAAATTAATGATATGATGATGGAAGGATGTTTGAAGAACTCATATGAAATTTTATC 179
QY 537 TCAAGTAACCTTATCCAGCCGAGATGAGTACACTTAAATGATGAGATGATGATCA 596
DB 180 TGAAGTAAGTAACTGATTAATTAATGAATGAATTTGTTTAACTGACAAATGTAAC 239
QY 597 TAAATCTCTTTGATGCGGAGACGGTACGTTTATCAAAATATTCAGCAATTTATCA 656
DB 240 TGAACCTGCTGACATGAAAGATTTGATCTGAAGAAAGCAACCTGAAACGCTGAAA 299
QY 657 TGAAGCAAGTGAATTCAGTGAATCAACACTTTTAAAGAGACCTGACTATGTTGAGA 716
DB 300 TGAAGAAAGTATGATGAGAGAGCTCATCAGAAAGAAATTTTACTGACAAATGATGA 359
QY 717 TAAAGGCCATTAATATCTTCTTTGAAGCAATATCTGAAACAAGATGTTATCAAG 776
DB 360 AGAATCAGGTGAAGTGATGATTTGATTAATGAAGAGGTGATTTTGAAGAACTAATCA 419

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:50:32 ; Search time 3489.14 Seconds
(without alignments)
11177.167 Million cell updates/sec

Title: US-09-986-682b-5

Perfect score: 2408
Sequence: 1 CCGGGAATAACTAGATTCC.....TACCAATAAATAATGTCATC 2408

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359.6	14.9	755	17	B49063 RPC11-4121
2	309.6	12.9	623	17	AQ284271
3	307.4	12.8	654	17	AG158197
4	296.8	12.3	687	17	AG145469
5	292.6	12.2	689	17	AG174787
6	290.8	12.1	653	17	AQ311782

7	286.6	11.9	688	17	AG165661	Pan trogl
8	284.4	11.8	656	17	AG158505	Pan trogl
c 9	283.4	11.8	678	17	AG019951	Homo sapi
c 10	271	11.3	711	17	AG001039	Homo sapi
11	267	11.1	669	17	AG156831	Pan trogl
12	264.8	11.0	381	17	AQ791204	HS_4555_B
c 13	258.4	10.7	682	17	AG020006	Homo sapi
14	250	10.4	934	9	AL575209	AL575209
15	248.6	10.3	683	17	AG185716	Pan trogl
16	233	9.7	666	17	AG167579	Pan trogl
17	228.6	9.5	622	17	AG160823	Pan trogl
18	226	9.4	508	17	AQ317303	RPC11-10
19	220.8	9.2	1101	17	CNS0005C	AL068761 Drosophila
20	217.4	9.0	687	17	AG145995	Pan trogl
21	215.8	9.0	488	17	AQ315317	AG155317 RPC11-10
22	215.2	8.9	661	17	AG019210	Homo sapi
23	215.2	8.9	678	17	AG020005	Homo sapi
24	208.2	8.6	694	17	AG145903	Pan trogl
25	205.4	8.5	694	17	AG169167	Pan trogl
26	203.8	8.5	698	17	AG165439	Pan trogl
27	203.6	8.5	682	17	AG165125	Pan trogl
28	199.6	8.3	674	17	AG165337	Pan trogl
29	199.6	8.3	686	17	AG177482	Pan trogl
30	193	8.0	601	17	P859L	AG177482 Pan trogl
31	192.8	8.0	698	17	AG180444	AG180444 Pan trogl
32	191.4	7.9	683	17	AG168295	AG168295 Pan trogl
33	189.4	7.9	705	17	AG147097	AG147097 Pan trogl
34	188.2	7.8	676	17	AG145781	AG145781 Pan trogl
35	186.6	7.7	385	17	B47667	B47667 RPC11-1A4
36	182.6	7.6	714	17	AG014172	AG014172 Homo sapi
37	180.6	7.5	613	17	AQ640465	AQ640465 927P1-13C
38	176.8	7.3	752	17	AG171011	Pan trogl
39	176.2	7.3	704	17	AG168501	Pan trogl
40	173.6	7.2	729	17	AG014170	Homo sapi
41	172.8	7.2	731	17	AG162359	Pan trogl
42	167.6	7.0	690	17	AG165865	Pan trogl
c 43	166.8	6.9	722	17	AG014192	Homo sapi
44	165.6	6.9	686	17	AG167003	Pan trogl
45	163.2	6.8	935	17	CNS0083C	AL052150 Drosophila

ALIGNMENTS

RESULT 1
LOCUS B49063
DEFINITION RPC11-4121.TV RPC1-11 Homo sapiens genomic clone RPC1-11-4121, DNA
ACCESSION B49063
VERSION B49063.1 GI:2601300
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 755)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden
,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter
,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPC1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

D _b	301	GCGGCGCAATGTTTTCCAAATGAAGCGCAATTGGTACTCTGTACACTGATTCACGCCGTTCC	360
O _y	1500	CAAAATGACGCGTGATGGAATTAACGACAAGAATGTTATATGCTAGGAGCCCGAAGCCGA	1558
D _b	361	AAAAATGACGATCATGGTATTAACTCAAAAGATTTTACATGGTTGGTTATGATCAA	420
O _y	1560	CTCCTTAAATGGCCACACCAACCGGATTAATGAATACGACTTGTATGAACATGAATCT	1619
D _b	421	CTCTNTAACCGCGCCTTACAAAGCCGCCTAACAAAACAGGGCTGTGCTGCGCAATGGGCTCT	480
O _y	1620	TGACCCTCGCTATCTMCACACACTACTACTCTCATTTGGGGTATCCCGCACCGCTGAAGTAA	1679
D _b	481	TGATTCCAACAGATGTGACATTCCTACTTCTCTCACTTCCGAGTGGCCGCAAGCCAAGGCTAA	540
O _y	1680	TAAATGTGACTCTACACAGTTATATGACGAATGAGGCTTCT	1720
D _b	541	CAATGTGGTTATCTCAACAGCTTACATGACACAAACGAGGGCTTCT	581

[illegible]

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library RPCI-43
Unpublished
2 (bases 1 to 654)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

COMMENT
Clones are derived from the chimpanzee BAC library RPT-43. This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS

Sequencing: T7
LIBRARY

Vector	:	pBACE3
R.Site 1	:	ECORI
R.Site 2	:	ECORI.

FEATURES	Location/Qualifiers
source	1. .654

BASE COUNT	ORIGIN
216 a	151 c 140 g 146 t 1 others

Query Match	12.8%	Score 307.4	DB 17	Length 654
Best Local Similarity	70.9%	Pred. No. 6e-69		
Matches 421; Conservative	0	Mismatches 172;	Indels 1	Gaps 1;

1128 TTCAGGTGTAACCATCTTCTTAAGAGAGACCCTCAGTA-TGTTGAAGATAAGGGCCATTAAT 1186

Db	22	TTCCGGGGACACCATATAGCTGAGAGACCCCTCACTACNCGTTGAAGACAAAGGCCATTAAT	81
OY	1187	ATCTTGTCTTTGAAGCGAATACTGGAAACAACAGATGGTTATCAAGGGCATCTTTTCA	1246
Db	82	ACCTTGTAATTCGAAAGCCACACGGGAGACGAAACGGATACCAAGGGAGAAATCTTTAT	141
OY	1247	ATATTAAGCTTACTATGGCGGAAGTGAAGCTTCTTCCAGATGAAHAAAATTAATACGC	1306
Db	142	TTAACAAAGCGCTACTACGGCGCGGACAGAACTTCTCCGTAAGAAAGCAGAAAGCTTC	201
OY	1307	TTCAAGTCCTAAAAAACAATTTGCTCTTTAGGAATGGATGGCATTTGGCATTTGAT	1366
Db	202	AGCAGAGCGCTAAAAAAGCGGATCTGTAGTTAGGAAACGGCGCCTCGGTAATAGAGT	261
OY	1367	TGGCCGATGACTATATACAGTGAAGAGTGTATGAACAACATTAGTCGATCAACACAGTAG	1426
Db	262	TAAATATATGATTAACATTCATTAAGAAAGTAATGAAGCCGCTGATCATCTTCAACACGGTAA	321
OY	1427	CAGATGAAGTGAACGGCCCAATATATTTAAATGAATTAATATGGTATCTATTACGG	1486
Db	322	CTGATGAATTCGACGGCGCGAATTTTTCAAAATGAACGGCAAAATGGTACTTGTTCACGT	381
OY	1487	ATTCAAGAGATCCAAAATGACGAGTGGATGGAATTAACGACAAAGATGTTATATTCAG	1546
Db	382	ATTACCGGGTTCAAAAATGACGATCGATGGTATTAACGACGATATTACATGGCTTG	441
OY	1547	GGCCCGGAGGACCTCCTTAATATGGCCACACACCCGATAAATGAACATGGACTTGTAT	1606
Db	442	GTTATGTATCAAACTCTTTAAACGGCCCTTAACAAGCCGCTGAACAACAAAGGGCTTGGC	501
OY	1607	TGAACATGAATCTTGAACCCCTGCTGATCTCACACACACTTACTCTCATTTGGCGATCCGC	1666
Db	502	TGCAAAATGGGCTTGATTCAAACAGCATGTGACATTCACCTTACTCTCACCTTGGCAGATGCCG	561
OY	1667	ACCCGTGAAGTAATATGTGGTACTACCAAGTTATATAGCAATAGAGGCTTCT	1720
Db	562	AAGCCAAAGGCAACAATGTGGTTATTCACAACACTACATGACAAACAGAGGCTTCT	615

RESULT 4	AG145469	687 bp	DNA	Linear	GSS 08-JAN-2002
LOCUS	AG145469				
DEFINITION	Pan troglodytes DNA, clone: RP43-006M11.T7, genomic survey				
ACCESSION	AG145469				
VERSION	AG145469.1				
KEYWORDS	GSS.				
SOURCE	Pan troglodytes male lymphocytes DNA, clone.lib:RPC1-43 Chimpanzee				

ORGANISM

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

AUTHORS	TITLE
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.	BAC end sequences of Library RPCI-43

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
Direct Submission
Tokoi, Y., Watanabe, H. and Sakaki, Y.
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
2 (bases 1 to 687)
unpublished

COMMENT¹

end was generated during the R&D process and may have higher chance of clone tracking errors.

Sequencing: T7
LIBRARY
Vector : pBACE3.6

R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. .687
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-006M11.17"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

BASE COUNT 239 a 148 c 147 g 153 t
ORIGIN

Query Match 12.3% Score 296.8; DB 17; Length 687;
Best Local Similarity 68.7%; Pred. No. 3.4e-66;
Matches 463; Conservative 0; Mismatches 187; Indels 24; Gaps 3;

```

OY 874 GAGTGTGAGTTCCTGCTTACTTAAACCAAGATGCGCAAGTCCGTTTATCTATACAT 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 30 GATGTGTCCTGCTTGCACACTTACATCTGACGCAAAATCTGTTTATCTACACTGAC 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 934 TACTGAGTAACTCTGAAGATGCTGAACCGGTGCTGTAACCAATCATTTCAACTGCT 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 90 TATTCGGTAACATTAC-----GGCAACAAAGCCTGCACACAGCG 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 994 CAAGTAACCTATCCACGCGGATCAGCTACCTTAAAGTGAAGTATCTGATCAT 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 132 CAGGTAATGCTCAAACTGATGAC--ACACTCAAAATCAACGAGTGAAGATGAC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1054 AAATGCTCTTGTGATGGGAGGCTACAGTTTCAAAATATTCAGATTTTCAT 1113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 AAAACGATTTTGA---CGGACGCAAAAACATATCAGAGCTTACAGATTTATCAT 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1114 GAAGCAAGTGTATTCAGTGATACCATATCTTAAAGACCTCAGTATGTTGAAGAT 1173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 246 GAAGCAATATATACCTCGCGACCAACCATACGCTGAGAGACCTCAGTATGTTGAAG 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1174 AAGGCCCAATATATCTGCTTAAAGCGATACTGACACAGATGTTATCAAGGC 1233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 306 AAAGGCCCAATATATCTGCTTAAAGCGATACTGACACAGATGTTATCAAGGC 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1234 GATCAGTCTTCAATATTAAGCTTACTATGCGGAGAGTGAAGTCTTCTCAGAGTAA 1293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 GAAGATCTTATTTAAACAAACGCTACGAGCGGCGGACAGACTTCTCCGTAAAGAA 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1294 AAAAATAACGCTTCAAGTCTTAAACAAATTTGCTTCTTAAAGCAATGCTGATTA 1353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 AGCCAGAGCTTCAAGCAGCGCTTAAACGCGATGCTGAGTTAGCGAAGCGCCCTC 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1354 GGCATGTTGAATGGCGGATGACATACAGTGAAGAAAGTGTATTAACCATTAAGTCCA 1413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 486 GGTATCATAGAGTTAAATATGATTACACATTTGAAGAAAGTAAAGCCCTGATCACT 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1414 TCAACAGATAGCAGATGAAGTGAAGCAAGCGCAATATTTAAATGAATTAATTAAGG 1473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 546 TCAACAGATAGCAGATGAAGTGAAGCAAGCGCAATATTTAAATGAATTAATTAAGG 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1474 TATCATTCACAGATTCAGAGATCCAAATAGACAGTATGAATTAACGACAAAGAT 1533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 606 TACTTGTCTCACTGATTCACGCGGTTCAAAATAGACGATGATGATTAATCAACAGAT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1534 GTTATATGCTAGG 1547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 666 ATTACATGCTTG 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5
AG174787 689 bp DNA linear GSS 09-JAN-2002
LOCUS AG174787
DEFINITION Pan troglodytes DNA, clone: RP43-045E23.17, genomic survey
sequence.
ACCESSION AG174787
VERSION AG174787.1 GI:16704467

KEYWORDS GSS.
SOURCE Pan troglodytes male lymphocytes DNA, clone:RPCI-43 Chimpanzee
Male BAC Library clone:RP43-045E23.17.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoh,Y., Matanabe,H. and Sakaki,Y.
TITLE BAC end sequences of library RPCI-43
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 689)
TITLE Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoh,Y., Matanabe,H. and Sakaki,Y.
JOURNAL Direct Submission
TITLE Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the Rad process and may have higher chance
of clone tracking errors.

PRIMERS

Sequencing: T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI

FEATURES
source
1. .689
Location/Qualifiers
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-045E23.17"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

BASE COUNT 240 a 151 c 143 g 154 t 1 others
ORIGIN

Query Match 12.2% Score 292.6; DB 17; Length 689;
Best Local Similarity 68.9%; Pred. No. 4.2e-65;
Matches 432; Conservative 0; Mismatches 189; Indels 6; Gaps 2;

```

OY 947 CTGAAGATGTGGAACCGGTGCTGCTGTAACCAATCATTTCAACGCTCAAGTAACTAT 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 CTGACTATTCGGTAACATATTAACGCAACAAACGCTGACACACGCCGAGTAATGTGT 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1007 CCCAGCGGATGACGCTACCTTAAAGTGAAGTATCTGATCATTAATCTGCTTTG 1066
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 CAAATCTGAT---GACACACTCAAAATCAACGAGAGTGAAGATCACAAAACGATTTTGG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1067 ATGGGAGAGAGGTACAGTTTATCAAAATATTGCAATTTATGCAATTAAGGCAAGTGA 1126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186 A---CGGAGACGCAAAACATATCAAGCTTCAAGCTTATGATGAAGCAATTAATA 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1127 TTTCAGGTGATTAACCATATTTAGAGACCTCAGTATGTTGAAGATTAAGGCCATTAAT 1186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 CATCCGGGACCAACCATATACGCTGAGAGACCTCAGTATGTTGAAGACAAAGCCATTAAT 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1187 ATCTTGTCTTGAAGCAATAGTGAACACAGATGTTATCAAGGCGATCAGTCTTTCA 1246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 ACCTTGTATTCGAAGCAACGCGGACAGAAACGAGATACCAAGGCGAAGAACTTTAT 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1247 ATATTAAGCTTACTATGCGGAGATGACGCTCTTCCAGAAATGAAGAAAAATTAACCTGC 1306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 363 TTAAACAAAGGCTACTACGCGCGGCGGACGAACCTTCTCCGTAAGAAAGCCAGAGCTTC 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1307 TTCAAGTCTTAACCAAAATTTGCTTGAAGATGTTGATGATGATGATGATGAT 1366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 423 AGCAGAGCGCTAAACAAACGCGATCTGAGTTAGGAACGCGCCCTCGGTATCATAGAGT 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1367 TGCGCGATGACTATACAGTGAAGAGTGTATGAACCATTAAGTGCATCAACACAGTAG 1426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      483 TAAATATGATTCACATTGAAAAAAGTAATGAAGCCGGTGCATCCTCAACACGGTAA 542
QY      1427 CAGATGAGTCGACACCGCCCAATATATTAAATGAATATATATGATCTTTTCACGG 1486
Db      543 CTGATGAAATCGAGCGCGCAATGTTTCAAAATGAACGGCAATGGTACTTGTCTACTG 602
QY      1487 ATTCAAGAGATCAAAATGAGCAGTGTGATTAACGCAAAAGATGTTTATATGCTAG 1546
Db      603 ATTACAGCGGTTCAAAAATGACGATGATGTATTAACTCAACAGATATTATCATGCTTG 662
QY      1547 GCGCGGAGCGACCTCTTAATGGCC 1573
Db      663 GTTATGTATCAACCTCTTAACCGGCC 689

RESULT 6
AO311782 653 bp DNA linear GSS 04-MAY-1999
LOCUS    RPC11-103A24.TV RPC1-11 Homo sapiens genomic clone RPC1-11-103A24,
DEFINITION
DNA sequence.
ACCESSION AO311782
VERSION    AO311782.1 GI:4043531
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 653)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE      Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbeetlir.org
            Clones are derived from the human BAC library RPC1-11. For BAC
            library availability, please contact Pieter de Jong
            (pieterdejong.med.bufileo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES
     source              1..653
                        Location/Qualifiers
                        /organism="Homo sapiens"
                        /db_xref="GDB:7539191"
                        /db_xref="taxon:9606"
                        /clone="RPC1-11-103A24"
                        /clone_lib="RPC1-11"
                        /sex="Male"
                        /cell_type="Lymphocytes"
                        /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
                        RPC11 Human Male BAC Library"
BASE COUNT 232 a 145 c 143 g 133 t
ORIGIN
Query Match 12.1%; Score 290.8; DB 17; Length 653;
Best local Similarity 68.0%; Pred. No. 1.2e-64;
Matches 460; Conservative 0; Mismatches 192; Indels 24; Gaps 3;

QY      787 ATTGACAGCTGAAAAATGCTGGAAGATATTGAAGATATGATAATTGTTCCAAT 846
Db      2 ATCGACAGCTGGAACACCGCGCCCTGTCTTAAAGACAGCATAGTTTCAGCGCAAC 61
QY      847 GATCCGATCTTAAATATCAACACAGAGGTGCTGAGTTTCTGCTACTTTAACCAAGAT 906
Db      62 GATCCGATCCTGAAAGATCAGACGAAGATGTCGCTTCGCAACCTTACATCTGAC 121

```

```

QY      907 GGGCAAGTCGGTTTATTTATACAGATTAAGTAATCTGAAGATGGTGAACGGT 966
Db      122 GGAATAATCCGTTTATTTACTACTGACTATTCGGTAACATTA----- 165
QY      967 GCTGTACCAAAATTCATTTCACTGCTCAAGTAAATCTTATCCAGCGGATCGCTACGA 1026
Db      166 --CGGCAACAAAGCCCTGACACAGCGGAGTAAATGTGCAAAATCTGAT---GACACA 220
QY      1027 CTTAAAGTCGATGAGTATCTGATTCATTAATTCCTCTTTGATGGGAGAGCGGTACAGT 1086
Db      221 CTCAAAATCAACGAGATGGAAGATCACAAAACGATTTTGA---CGAGACGGAAGAAACA 277
QY      1087 TATCAAAATATATACGAATTTATGATGTAAGGACGATGTTTACAGTATTAACCTACT 1146
Db      278 TATCAGACGTTTATGAGATTTATGATGTAAGGCAATTAATACATCGGCAACCATACG 337
QY      1147 TTAAGAGACCCCTCACTATGTTGAAGATTAAGGGCCATTAATATCTGTTTGAAGCGAAT 1206
Db      338 CTGAGAGACCCCTCACTATGTTGAAGATTAAGGGCCATTAATATCTGTTTGAAGCGAAT 397
QY      1207 ACTGAAACACAGATGTTATCAAGCGGATCACTTTCAATATAAAGCTTACTATGGC 1266
Db      398 ACGGAAACAGAAACGAGTACCAAGCGGAAGATCTTTATTTAAACAAAGCGTACTACGCC 457
QY      1267 GGAAGTGACGCTCTCTCCAGATGAAAAAATAAATCTGTTCAAGTCTCTTAAAAACAA 1326
Db      458 GGGGACACGACGATCTTCCGTAAGAAAGCCAAAGCTTACGAGCGCTTAAAAACCG 517
QY      1327 ATTGCTTCTTTTACGAGATGTTGATGATGATGATGATGATGATGATGATGATGATG 1386
Db      518 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 577
QY      1387 AAAAGTGTATGAACCATTAATGTCGATCAACACAGTACAGATGATGATGATGATGATG 1446
Db      578 AAAAAGTATGAAGCGCGCTGATCACTTATACACGATGATGATGATGATGATGATGATGATG 637
QY      1447 AATATATTAAATGA 1462
Db      638 AATGTTTCAAAATGA 653

RESULT 7
AG165661 688 bp DNA linear GSS 09-JAN-2002
LOCUS    Pan troglodytes DNA, clone: RP43-033D03.T7, genomic survey
DEFINITION
sequence.
ACCESSION AG165661
VERSION    AG165661.1 GI:16695339
KEYWORDS   GSS.
SOURCE     Pan troglodytes male lymphocytes DNA, clone_lib:RPC1-43 Chimpanzee
            Male BAC Library clone:RP43-033D03.T7.
            Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; catarrhini; Homnidae; Pan.
ORGANISM   Pan troglodytes
REFERENCE 1
AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library RPC1-43
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 688)
AUTHORS   Fujiyama,A., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suenho-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpses@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT    Clones are derived from the chimpanzee BAC library RPC1-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
            PRIMERS

```

```

Sequencing: T7
LIBRARY
Vector : PBAC3.6
R.Site 1 : EcorI
R.Site 2 : EcorI.
Location/Qualifiers
1. 688
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-033D03.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      227 a      160 c      154 g      147 t
ORIGIN

Query Match      11.8%; Score 286.6; DB 17; Length 688;
Best Local Similarity 68.4%; Pred. No. 1.5e-63;
Matches 451; Conservative 0; Mismatches 184; Indels 24; Gaps 3;

QY 627 AATAGATTAGATGTATGATAGTGGCCATGCAAAACGTGATGCTGCGCAAA 686
    || || || || || || || || || || || || || || || || || || ||
Db 53 AAAAGAGCTGATGTGGGACAGCTGGCCGCTGCACAAAGCTGACGGAACACTAGCTGA 112

QY 687 TTATCATGATATACATGCTCTCCGCTTATGACAGCTGACCCAAAAACAGTATGATAC 746
    || || || || || || || || || || || || || || || || || || ||
Db 113 ATACAAAGGCTATACAGTGTGTGCTTGTGCGGGAAGCCGAAAGAGCTGATGACAC 172

QY 747 TCACCTTCATTTATCATCAAAAGTCGGGTGATACATGATGACAGCTGGAATAATGC 806
    || || || || || || || || || || || || || || || || || || ||
Db 173 ATCAATCTACATGTTTATCAAAAGTCGGGACAACTCAATGACAGCTGGAATAATGC 232

QY 807 TGAAGAGTATTTGAGATATGATTAATTTGTTCCAAATGATCCGATCTTAAATATCA 866
    || || || || || || || || || || || || || || || || || || ||
Db 233 GGGCCGCTGTTTAAAGACAGCATTAAGTTCGACGCCAACAGATCCGATCCTGAAGATCA 292

QY 867 AACACAGAGAGTGTCTGAGTCTCTACTTTAACCAAGATGSCCAAGTCCGTTATTCTA 926
    || || || || || || || || || || || || || || || || || || ||
Db 293 GACGCAAGATGTGTCGGTCTGCAACCTTACATCTGACGGAATAATCCGTTATTCTTA 352

QY 927 TACAGATTACTGAGTATCTGTAAGATGTGTGACCGGTCGTTACCAAAATCAATTTC 986
    || || || || || || || || || || || || || || || || || || ||
Db 353 CACTGACTATTCGGGTAACATTA-----CGCAAAACAAAGCCCTGAC 394

QY 987 AACTGCTCAAGTAACCTATCCAGCGGATGACGCTACACTTAAAGTCGATGAGATATC 1046
    || || || || || || || || || || || || || || || || || || ||
Db 395 AACAGCGCAGTAAATGTGTCAAAATCTGAT---GACACACTCAAAATCAACGAGTGA 451

QY 1047 TGAATCAATATCTGCTTGATGGCGGAGACGATACAGTTATCAAAATTTAGCAAT 1106
    || || || || || || || || || || || || || || || || || || ||
Db 452 AGATCAACAAACGATTTTGA---CGAGACGGAATAATATCAGAACGTTGACGAGTT 508

QY 1107 TATCATGTAAGGCAAGTGTATTCAGGTGATTAACCACTTAAAGAGACCTCATATGT 1166
    || || || || || || || || || || || || || || || || || || ||
Db 509 TATCATGTAAGGCAATATATCATCTCCGCGACCAACCATACGCTGAGAGACCTCATATGT 568

QY 1167 TGAAGTAAGGCGCATTAATATCTGTCTTTGAAGCAATAGTGAACAACAGATGTTA 1226
    || || || || || || || || || || || || || || || || || || ||
Db 569 TGAAGCAAAAGGCAATAATATCTGTATCGAAGCAACACGGAACAAGAAAACGATTA 628

QY 1227 TCAAGGCAATCACTTTCATATAAAGCTTACTATGGGGAGAGTACGCTCTTTC 1285
    || || || || || || || || || || || || || || || || || || ||
Db 629 CCAAGGCAAGATCTTATTTTACAAAGGCTACTACGGGGGCGGACGAACTTCTTCC 687

RESULT 8
AG158505      656 bp      DNA      linear      GSS 09-JAN-2002
LOCUS         Pan troglodytes DNA, clone: RP43-023M01.T7, genomic survey
ACCESSION    AG158505
VERSION      AG158505.1  GI:16688183
KEYWORDS     GSS.

```

```

SOURCE
ORGANISM
Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
Male BAC Library clone:RP43-023M01.T7.
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.

REFERENCE
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library RPCI-43
Unpublished
2 (bases 1 to 656)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbases@sc.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS

FEATURES
source
1. 656
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-023M01.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      229 a      144 c      140 g      143 t
ORIGIN

Query Match      11.8%; Score 284.4; DB 17; Length 656;
Best Local Similarity 69.6%; Pred. No. 5.7e-63;
Matches 416; Conservative 0; Mismatches 176; Indels 6; Gaps 2;

QY 947 CTGAAAGTGTGGACCGGTGCTGTACCAATCAATTTCACTGCTCAAGTAACCTTAT 1006
    || || || || || || || || || || || || || || || || || || ||
Db 65 CTGACTATTCGGTAAACATTAACGCAAAAGCCTGACACAGCGCAGTAATATGCT 124

QY 1007 CCCAGCGGATGAGCTACACTTAAGTGTGAGATATCTGATCATTAATCTGCTTTG 1066
    || || || || || || || || || || || || || || || || || || ||
Db 125 CAAATCTGAT---GACACACTCAAAATCAACGAGTGAAGATCAAAACGATTTTGG 181

QY 1067 ATGGCGAGACGCTACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGCAAGTGA 1126
    || || || || || || || || || || || || || || || || || || ||
Db 182 A---CGAGACGGAATAACATATCAACAGCTTCAGCAGTTTATCATGAAGCAATTTATA 238

QY 1127 TTTACAGGTATTAACCACTTATTAAGAGACCTCATCTTTGAAGTAAGGCCATTAAT 1186
    || || || || || || || || || || || || || || || || || || ||
Db 239 CATCCGCGACAACCACTATAGCTGAGAGACCTCATCACTTGAAGCAAAAGGCCATTAAT 298

QY 1187 ATCTGCTTTGAAGCAATCTGAGAACAGATGCTGTAATGAAGCGCATGCTTTCA 1246
    || || || || || || || || || || || || || || || || || || ||
Db 299 ACCTGTATTCGAAGCAACAGCAAGCAAGAAACGATACCAAGCGCAAGATCTTTAT 358

QY 1247 ATATAAAGCTTACTATGGCGAGAGAGAGTCTTTCCTCAGATGAAGAAAATAAATCAG 1306
    || || || || || || || || || || || || || || || || || || ||
Db 359 TTAACAAGCGTACTACGGCGGCGGCGCAGCAACTTCTCCGTAAAGAAACCAAGACTTC 418

QY 1307 TTCAAAGCTCTAAACAAATTTGCTTTTAAAGCAATGTCATTAAGCATTTGTTGAAT 1366
    || || || || || || || || || || || || || || || || || || ||
Db 419 AGCAGAGCGCTAAACAAAGCGATGCTGATGAGCAAGCGGCCCTCGTATCATATGAGAT 478

QY 1367 TGGCCGATGACTATACAGTGAAGAAAGTGTATGAACCAATTAAGTCGATCAACACAGT 1426
    || || || || || || || || || || || || || || || || || || ||

```



```

Db 590 GACGGAAAAATCCGTTTATCTACACTGACTATTCGGTAAACAT----- 546
OY 964 GTGTGTGTGTAACCAATCATTTCACTGCTCAAGTAATATTCACCGGATGCACCT 1023
Db 545 --TAGCGCCAAACAAAGCCTGACACAGCAGATGTAATGTCTCAAAATCTGAT---GAC 491
OY 1024 ACATTTAAAGTCATGATGATATCTGATCAATATCTGTTTGAATGCGGAGAGCGTACA 1083
Db 490 ACACCAAAATCAACAGAGTGAAGATCACAACAAAGATTTTGA---CGGAGACGGAAAA 434
OY 1084 GTTATCAAAATATTCACCAATTTATCATGATGAAGCAGATTCAGTGTGATACCAT 1143
Db 433 ACATATCGAAAGCTTCACAGATTTATCGATGAAGGCAATTAATACATCGGAGACACCAT 374
OY 1144 ACTTTAAGAGACCCCTCATGTATGTAAGATGAAGCCCAATATCTGTTTGAACGC 1203
Db 373 ACAGTCAGAGACCCCTCATGTATGTAAGATGAAGCCCAATATCTGTTTGAACGC 314
OY 1204 AATACTGACACACAGATGTTATCAAGGCAATCAGTCTTTCATATTAAGCTTACTAT 1263
Db 313 AACACGGGACACAGAAAGCGATACCAAGCGAAGATCTTTATTTAACAAAGCGTATAC 254
OY 1264 GCGGCAATGACGCTTCTTCAGATGAAAAAATAAAGCTTCAAGTCTTAATAAA 1323
Db 253 GCGGCGGCGACAGACTTCTTCGTAAGAAAGCCAGAACTTCAGACGAGCTTAATAAA 194
OY 1324 CAATTTGCTCTTAAAGCAATGATGTCATTAAGCATTTGTTGAATGGCCGATGACTATACA 1383
Db 193 CGCGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 134
OY 1384 GTGAAAGTGTATGAAACCAATTAAGTCGATCAACACAGATGATGATGATGATGATGATGAT 1443
Db 133 TTGAAAAAGTAAATGAAAGCCGCTGATCACTTCAACACGATGATGATGATGATGATGATGAT 74
OY 1444 GCCAATATATTTAAATGATTAATTAATGATA 1475
Db 73 GCGAATGTTTCAAAATGACAGCGCAATGTGA 42

RESULT 11
AG156831 669 bp DNA Linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RP43-021J04.T7, genomic survey
DEFINITION
ACCESSION AG156831.1 GI:16686509
VERSION AG156831.1
KEYWORDS
SOURCE Pan troglodytes male lymphocytes DNA, clone:lib:RP43-43 Chimpanzee
Male BAC Library clone:RP43-021J04.T7.
ORGANISM
REFERENCE
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library RP43-43
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 669)
TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Tokoki, Y., Watanabe, H. and Sakaki, Y.
JOURNAL Direct Submission
AUTHORS Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RP43-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T7
LIBRARY Vector : pBACe3.6

```

```

FEATURES
Source
R.site 1 : Ecoli
R.site 2 : Ecoli.
Location/Qualifiers
1..669
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-021J04.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP43-43 Chimpanzee Male BAC Library"
BASE COUNT 227 a 153 c 143 g 146 t
ORIGIN
Query Match 11.1%; Score 267; DB 17; Length 669;
Best Local Similarity 65.5%; Pred. No. 1.9e-58;
Matches 390; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

OY 1205 ATACTGAAACACAGATGTTATCAAGGCGATCAGTCTTTCATATTAATTAAGCTTATG 1264
Db 34 ACACGGGACACAGAAAGCGATACCAAGCGAAGATCTTTATTTAACAAAGCGTATACG 93
OY 1265 GCGGAGTGAAGCTCTCTTCAGATGAAAAAATAAAGCTTCAAGTCTTAATAAAAC 1324
Db 94 GCGGCGGCGACAGACTTCTTCGTAAGAAAGCCAGAACTTCAGACGAGCGCTTAATAAAAC 153
OY 1325 AAATTTGCTCTTAAAGCAATGATGTCATTAAGCATTTGTTGAATGGCCGATGACTATACA 1384
Db 154 GCGATGCTGATTTAGCGAAGCGCCCTCGGTATCATAGATTAATTAATGATTAACAT 213
OY 1385 TGAAGAAGTGTATGAACCAATTAAGTCGATCAACACAGTGAAGATGAAGTGAACGCG 1444
Db 214 TGAAGAAAGTATGAACCGCTGATCACTTCAACACGATGAAGTGAAGTGAAGTGAAGTGAAG 273
OY 1445 CCATATATTTAAATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504
Db 274 GCAATGTTTCAAAATGACGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 333
OY 1505 TGACAGTGAATGAATTAACGACCAAGATGTTTATATGATGATGATGATGATGATGATGATGAT 1564
Db 334 TGACGATGATGATGATGATTAATTAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 393
OY 1565 TAAATGGCCCAACACCCGATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1624
Db 394 TAAACGGCCCTTAACAGCCGCTGAACAAAGAGGCTTGACGCAATGATGATGATGATGATGATGATGAT 453
OY 1625 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1684
Db 454 CAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 513
OY 1685 TGGTACTGACAAAGTATATGAGAGGCTTCTATCCAGAAACATCACTTCCACCTGC 1744
Db 514 TGGTTATACACAGCTACATGACAAAGAGGCTTCTGAGAGATTAAGCAACATTTG 573
OY 1745 GGGCAAGCTTGGGCTTAATTTAAAGGCTGACACATCTGGAGGAGAAATATG 1799
Db 574 GCGCAACCTTCTTAATGACATCAAGCAATTAACATCCGTTGCAAAAAACAG 628

RESULT 12
LOCUS AO791204 581 bp DNA Linear GSS 03-AUG-1999
DEFINITION HS-4555_B1_A10.T7A CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-4555 Col-19 Row-B, DNA sequence.
ACCESSION AO791204
VERSION AO791204.1 GI:5698916
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Maitiras, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and

```


Query Match	10.4%	Score 250;	DB 9;	Length 934;
Best Local Similarity	64.8%;	Pred. No. 5.1e-54;		
Matches 367;	Conservative 2;	Mismatches 197;	Indels 0;	Gaps 0;
1234 GATCAGCTTTTCATTAATAATGCTTACTATGTCGGGAGCTGACGCTCTTTCCAGATGAA 1233				
23 GAAGATCTTTATTTAATAACAAAGCTACTACGCGCGGACAGAACTTCTCCGTAAGAA 82				
1294 AAAATAACGCTCTCAAGAGTCGTAAAAACAATTGCTCTTGAAGCAATGTCATTA 1353				
83 AGCAGAAAGCTTCACAGACGCGCTRAAAAAACGCGATGCTGAGTTAGCGAACGCGCCCTC 142				
1354 GGCATTGTTGAATTGGCCGATGACTATACAGTGAAGAAAGCTTTATGAACCAATTAAGTCGA 1413				
143 GGTATCATAGAGTTAATAATTAATGATTACACATTGAAAAAGTAATGAACCCGCTGATCACT 202				
1414 TCAAAACACAGTAGCAGATGAAGTGGAAACGCCCATATATTAAAAATGAATTAATTAATG 1473				
203 TCAACACAGTACTAGTAAGTAATGAGGCGCGGCAATGTTTCAAAATTAAGGCAAAATGG 262				
1474 TATCATATCCAGATATCAAGAGATCCAAATAGCAAGTGAAGTGAATTAAGCAACAAAGAT 1533				
263 TACTTGTTCACTGATTCACGCGGCTTCAAAAATGACATCGATGCTATTAACTCAAAAGAT 322				

QY	153	GTTTATTTGCTAGGCGCCGGAGGCGGACTCCCTTAATGGCCACACACCGGATTAATGA	1593
Db	323	ATTTCATAGCTGGTTATGTATCAAACTCTTTAACCGGCCCTTTCAAAAAMAAAAA	382
QY	1594	ACTGAGCTTGATATGAACATGAATCTTGACCCTGCTCATCTCACACACACTTACTTCAT	1653
Db	383	AAGAGGGCTTGTCGTCGCAAAATGGGCTTGATGTCACCAAGCATGTGACATTCACCTACTTCAC	442
QY	1654	TGGGGTATCCCGCACCTGAAAGGTAAATATGTGTACTCACAACTTATATGACAAATGA	1713
Db	443	TTCCGACTGCGCGCAAGCCAAAGGCAACAAATGCTGTATCATCAACACTCATGACAAACGA	502
QY	1714	GGCTTCATCCAGAACATCACTCTCACCTGGGGGACAAAGCTTGGGGTATATTAAGG	1773
Db	503	GGCTTCTTCGAGGTAAAGGCAACACTTTGGCCGCAAGCTTCTTAATGAACATCAAAAGC	562
QY	1774	TCTGACACATCTGGAGGAGAAATATG	1799
Db	563	AATAAACATCTCGTTGTCAAAACAG	588
RESULT 15			
AG185716			
LOCUS	AG185716	696 bp	DNA
DEFINITION	Pan troglodytes DNA, clone: RP43-059006.T7, genomic survey		
ACCESSION	AG185716		
VERSION	AG185716.1	GI:16715396	
KEYWORDS	GSS.		
SOURCE	Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee		
ORGANISM	Pan troglodytes		
REFERENCE	1		
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,		
TITLE	Totoki, Y., Watanabe, H. and Sakaki, Y.		
JOURNAL	BAC end sequences of Library RPCI-43		
REFERENCE	2 (bases 1 to 696)		
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,		
TITLE	Totoki, Y., Watanabe, H. and Sakaki, Y.		
JOURNAL	Direct Submission		
COMMENT	Submitted (02-AUG-2001) Aao Fujiyama, The Institute of Physical		
	and Chemical Research (RIKEN), Genomic Sciences Center (GSC);		
	1-7-22 Shehito-Chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan		
	(E-mail:chimbesc@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/		
	Tel:81-45-503-9111, Fax:81-45-503-9170)		
	Clones are derived from the chimpanzee BAC library RPCI-43 This BAC		
	end was generated during the R&D process and may have higher chance		
	of clone tracking errors.		
	PRIMERS		
	Sequencing: T7		
	LIBRARY		
	Vector : pBAC3.6		
	R.Site 1 : EcoRI		
	R.Site 2 : EcoRI		
	Location/Qualifiers		
	1..696		
	/organism="Pan troglodytes"		
	/db_xref="taxon:9598"		
	/clone="RP43-059006.T7"		
	/sex="male"		
	/cell_type="lymphocytes"		
	/clone_lib="RPCI-43 Chimpanzee Male BAC Library"		
	237 a 162 c 148 g 146 t 3 others		
BASE COUNT			
ORIGIN			
Query Match	10.3%	Score 248.6;	DB 17; Length 696;
Best Local Similarity	64.6%;	Pred. No. 1.2e-53;	
Matches 459; Conservative	0; Mismatches 200;	Indels 51; Gaps 4;	
473	ACAAGAAGACTATGTTTGGCCATATATACACGCGCTGACATCTCTAAATTCACAGAC	532	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 18, 2003, 07:25:42 ; Search time 365.061 Seconds
(without alignments)
9679.375 Million cell updates/sec

Title: US-09-986-682B-5
Perfect score: 2408
Sequence: 1 CGGGGAAATAGTAGATTCC.....TACCAATATAATGATGATC 2408

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085033

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2408	100.0	2408	10 US-09-986-682B-5	Sequence 5, Appl1
2	1365	56.7	1365	10 US-09-986-682B-5	Sequence 4, Appl1
3	627	26.0	1449	10 US-09-974-300-2644	Sequence 2644, Ap
4	608.8	25.3	4938	9 US-10-055-794-1	Sequence 1, Appl1
5	608.8	25.3	6525	9 US-10-055-794-3	Sequence 3, Appl1
6	607.2	25.2	10317	10 US-09-849-866-1	Sequence 1, Appl1
7	580.6	24.1	2253	10 US-09-993-292A-19	Sequence 19, Appl1
8	580.6	24.1	8908	10 US-09-993-292A-18	Sequence 18, Appl1
9	135.6	5.6	368	10 US-09-864-761-21746	Sequence 21746, A
10	110.4	4.6	882	10 US-09-974-300-735	Sequence 735, App
11	93	3.9	1146	10 US-09-778-844-136	Sequence 136, App
12	80.2	3.3	376	10 US-09-864-761-5018	Sequence 5018, Ap
13	61.2	2.5	4634	10 US-09-995-587A-10	Sequence 10, Appl1
14	52	2.2	839	9 US-10-198-846-5829	Sequence 5829, Ap
15	52	2.2	2367	10 US-09-995-587A-2	Sequence 2, Appl1
16	52	2.2	2394	10 US-09-995-587A-3	Sequence 3, Appl1
17	52	2.2	2592	10 US-09-995-587A-4	Sequence 4, Appl1
18	47	2.0	931	9 US-10-198-846-5362	Sequence 5362, Ap
19	44	1.8	58985	9 US-09-901-152-3	Sequence 3, Appl1

20	41.8	1.7	5301	7 US-08-781-986A-443	Sequence 443, App
21	39	1.6	394	10 US-09-878-574-3784	Sequence 3784, Ap
22	38.6	1.6	921	10 US-09-815-242-4588	Sequence 4588, Ap
23	38.6	1.6	957	10 US-09-815-242-8509	Sequence 8509, Ap
24	38.4	1.6	5086	9 US-09-764-881-9794	Sequence 9794, Ap
25	38	1.6	314	10 US-09-833-381-190	Sequence 190, App
26	38	1.6	504	9 US-10-074-095-200	Sequence 200, App
27	38	1.6	548	10 US-09-764-860-200	Sequence 200, App
28	38	1.6	548	10 US-09-974-300-3392	Sequence 3392, Ap
29	38	1.6	1257	9 US-10-074-095-996	Sequence 996, App
30	38	1.6	1257	10 US-09-764-860-996	Sequence 996, App
31	37.8	1.6	544	9 US-10-198-846-13090	Sequence 13090, A
32	37.8	1.6	852	9 US-10-198-846-6037	Sequence 6037, Ap
33	37.8	1.6	3831	12 US-10-051-952-5	Sequence 5, Appl1
34	37.4	1.6	8761	9 US-10-098-841-11	Sequence 11, Appl1
35	37.4	1.6	10758	12 US-10-044-090-61	Sequence 61, Appl1
36	37.4	1.6	1830121	9 US-10-329-960-1	Sequence 1, Appl1
37	37.2	1.5	606	9 US-10-106-698-208	Sequence 208, App
38	37.2	1.5	606	10 US-09-925-301-120	Sequence 120, App
39	37.2	1.5	3996	9 US-10-087-464-42	Sequence 42, Appl1
40	37	1.5	1799	7 US-08-781-986A-447	Sequence 447, App
41	36.8	1.5	7153	9 US-10-239-676-69	Sequence 69, Appl1
42	36.6	1.5	7101	9 US-10-011-366-9	Sequence 9, Appl1
43	36.4	1.5	756	9 US-10-244-292-6	Sequence 6, Appl1
44	36.4	1.5	17142	9 US-10-239-676-305	Sequence 205, Appl1
45	36.2	1.5	715	9 US-10-198-846-336	Sequence 336, App

ALIGNMENTS

RESULT 1
US-09-986-682B-5
Sequence 5, Application US/09986682B
Patent No. US20020115182A1
GENERAL INFORMATION:
APPLICANT: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU
KENKYUJO
TITLE OF INVENTION: POLYPEPTIDE HAVING - FRUCTOFURANOSIDASE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/986, 682B
FILING DATE: 06-NO. US20020115182A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 170, 630/1996
FILING DATE: 10-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25, 618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2408 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus sp.
INDIVIDUAL ISOLATE: V230 (FERM BP-5054)
FEATURE:
NAME/KEY: signal peptide
LOCATION: 1..360
IDENTIFICATION METHOD: E
SEQUENCE DESCRIPTION: SEQ ID NO: 5
US-09-986-682b-5

Query Match 100.0%; Score 2408; DB 10; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGGGGAAATATCTGATTCATTCATTCGACAGCTTCCCACTGGTGTAGAGAGAGTTCG 60
DB 1 CGGGGAAATATCTGATTCATTCATTCGACAGCTTCCCACTGGTGTAGAGAGAGTTCG 60
OY 61 GACTGCCAATGCAGCTGTGCGTAAGAAACAGCTTACTCATGAGCAATTACTAGAGAT 120
DB 61 GACTGCCAATGCAGCTGTGCGTAAGAAACAGCTTACTCATGAGCAATTACTAGAGAT 120
OY 121 TTCAAAATCTCGGATTAAGCCCAAGTCCACTTTGAATTAACCTTTGAGCCTGTGTG 180
DB 121 TTCAAAATCTCGGATTAAGCCCAAGTCCACTTTGAATTAACCTTTGAGCCTGTGTG 180
OY 181 GGGGCTTTTGTATTTTATTTCACTGCAAGTGTCCATCCCTATATCAATTTAA 240
DB 181 GGGGCTTTTGTATTTTATTTCACTGCAAGTGTCCATCCCTATATCAATTTAA 240
OY 241 GACGAAATTTCAATCATCATGCCATCCCAATTAACCTGCTCTCTATCTTTTAA 300
DB 241 GACGAAATTTCAATCATCATGCCATCCCAATTAACCTGCTCTCTATCTTTTAA 300
OY 301 TTAATAAACAATTCAGAGCTTTCTTATCAATTCATATATCCAGAGAGAGAGC 360
DB 301 TTAATAAACAATTCAGAGCTTTCTTATCAATTCATATATCCAGAGAGAGAGC 360
OY 361 ATGAACCTTCAAAAGATTGGCGAAAAAAGACAGCTGCCGTAACCTTCAGAGCTTATTTA 420
DB 361 ATGAACCTTCAAAAGATTGGCGAAAAAAGACAGCTGCCGTAACCTTCAGAGCTTATTTA 420
OY 421 GTAGAGCGGAGCGACCGCATATTTTTCGCGAGCAATGAACAGCGGGGACTACAGGAA 480
DB 421 GTAGAGCGGAGCGACCGCATATTTTTCGCGAGCAATGAACAGCGGGGACTACAGGAA 480
OY 481 GACTATGTTTGGCCATATTAACAGCGCGCTGACATGCTAAAAATTCACGAGCAACAAAC 540
DB 481 GACTATGTTTGGCCATATTAACAGCGCGCTGACATGCTAAAAATTCACGAGCAACAAAC 540
OY 541 AGTCCTCAATTTAAAGTCCCTCAATTCATGCAATGAGCAATCAAAAATTTGATTCGCA 600
DB 541 AGTCCTCAATTTAAAGTCCCTCAATTCATGCAATGAGCAATCAAAAATTTGATTCGCA 600
OY 601 AAAGGATATGAATAGTCAAGCACTTAATAGATTAGATGATGAGGATAGCTGGCCACTG 660
DB 601 AAAGGATATGAATAGTCAAGCACTTAATAGATTAGATGATGAGGATAGCTGGCCACTG 660
OY 661 CAAAACGCTGATGAGTACGCGGCAAAATTAATCATGATATCAATCGTCTCGCTTTTACA 720
DB 661 CAAAACGCTGATGAGTACGCGGCAAAATTAATCATGATATCAATCGTCTCGCTTTTACA 720
OY 721 GGTGACCCAAAACAGTATGATATCTCACTTATTTCTATCAAAAAGTGGGTAT 780
DB 721 GGTGACCCAAAACAGTATGATATCTCACTTATTTCTATCAAAAAGTGGGTAT 780
OY 781 ACATGATATGACAGCTGGAAGAAATGCTGAAGAGATTTGAAGATATGATGATTTGTT 840
DB 781 ACATGATATGACAGCTGGAAGAAATGCTGAAGAGATTTGAAGATATGATGATTTGTT 840
OY 841 CCAATGATCGGATCTTAATATCAACACAGAGAGTGTGAGTTCGTACTTTAAC 900
DB 841 CCAATGATCGGATCTTAATATCAACACAGAGAGTGTGAGTTCGTACTTTAAC 900

DB 841 CCAATGATCGGATCTTAATATCAACACAGAGAGTGTGAGTTCGTACTTTAAC 900
OY 901 AAAGATGGCCCAAGTCCGTTTATTTCTATACATATTCAGAGTATCTCTGAAGATGGTGA 960
DB 901 AAAGATGGCCCAAGTCCGTTTATTTCTATACATATTCAGAGTATCTCTGAAGATGGTGA 960
OY 961 ACCGCTGTGTAAACCAATATCTCAACTGCTCAAGTAACTATATCCAGCCGATGCA 1020
DB 961 ACCGCTGTGTAAACCAATATCTCAACTGCTCAAGTAACTATATCCAGCCGATGCA 1020
OY 1021 GCTACACTTAAAGTGCATGAGTATCTGATCAATATCTGTTGATGGCGAGACGT 1080
DB 1021 GCTACACTTAAAGTGCATGAGTATCTGATCAATATCTGTTGATGGCGAGACGT 1080
OY 1081 ACAGTTTATCAAAATATTCACAGCAATTAATCAAGCAAGTGTGATTCAGGTGATAC 1140
DB 1081 ACAGTTTATCAAAATATTCACAGCAATTAATCAAGCAAGTGTGATTCAGGTGATAC 1140
OY 1141 CATACCTTAAAGAGACCTCACTATGTAAGATTAAGGCGCTAATATCTGTTGAA 1200
DB 1141 CATACCTTAAAGAGACCTCACTATGTAAGATTAAGGCGCTAATATCTGTTGAA 1200
OY 1201 GCGAATATGAGACACAGATGTTATCAAGCGCATAGTCTTCAATTAATTAAGCTTAC 1260
DB 1201 GCGAATATGAGACACAGATGTTATCAAGCGCATAGTCTTCAATTAATTAAGCTTAC 1260
OY 1261 TATGCGGAAATGAGCTGCTTCTTCAGATGAATAAATTAACGCTTCAAAAGCTTAA 1320
DB 1261 TATGCGGAAATGAGCTGCTTCTTCAGATGAATAAATTAACGCTTCAAAAGCTTAA 1320
OY 1321 AAACAAATGCTTCTTTAGCGAATGTCATTAAGCATTTGTTGAATTTGGCCGATGAT 1380
DB 1321 AAACAAATGCTTCTTTAGCGAATGTCATTAAGCATTTGTTGAATTTGGCCGATGAT 1380
OY 1381 ACAATGAAATGTTATGAACCAATAGTCCATCAAAACACAGTATGAAGTGA 1440
DB 1381 ACAATGAAATGTTATGAACCAATAGTCCATCAAAACACAGTATGAAGTGA 1440
OY 1441 CGGCGCAATATATTAATGAATGAATGAATGATGATTCACAGATTCAGAGAGTTC 1500
DB 1441 CGGCGCAATATATTAATGAATGAATGAATGATGATTCACAGATTCAGAGAGTTC 1500
OY 1501 AAATGACGATGATGAATTAACGACAAAGATGTTATATGTAAGGCGCGAGCGAC 1560
DB 1501 AAATGACGATGATGAATTAACGACAAAGATGTTATATGTAAGGCGCGAGCGAC 1560
OY 1561 TCCTTAATGAGCCCAACACACACCCGATTAATGAATGATGATGATGAATGAT 1620
DB 1561 TCCTTAATGAGCCCAACACACACCCGATTAATGAATGATGATGATGAATGAT 1620
OY 1621 GACCTGCTGATGTCACACACACTTACTGATGCGGTATCCGCAACCTGAAGATAT 1680
DB 1621 GACCTGCTGATGTCACACACACTTACTGATGCGGTATCCGCAACCTGAAGATAT 1680
OY 1681 AATGTGATCTACACAGTATATAGCAATAGAGGCTTATTCAGAACATCACTCTAC 1740
DB 1681 AATGTGATCTACACAGTATATAGCAATAGAGGCTTATTCAGAACATCACTCTAC 1740
OY 1741 CTGCGGGAACAGCTGGGGTTAATATTAAGGCTGACACATCTGGAAGGAATAATAGT 1800
DB 1741 CTGCGGGAACAGCTGGGGTTAATATTAAGGCTGACACATCTGGAAGGAATAATAGT 1800
OY 1801 TCCGAGACAGGACATTCCTATAGCATATCTCCCATTAATAAAGATGTCACCTGCA 1860
DB 1801 TCCGAGACAGGACATTCCTATAGCATATCTCCCATTAATAAAGATGTCACCTGCA 1860
OY 1861 GATCTACCTTTTCCCAACCTTTTAAAAATGAAGGTTTGAACACCTTTAGAA 1920
DB 1861 GATCTACCTTTTCCCAACCTTTTAAAAATGAAGGTTTGAACACCTTTAGAA 1920
OY 1921 GAAAAAAGATCTTTGGCGCGGCAATTAACCCATTCGCAAGAGGCTACCCCAATAAAA 1980
DB 1921 GAAAAAAGATCTTTGGCGCGGCAATTAACCCATTCGCAAGAGGCTACCCCAATAAAA 1980

QY 1981 AGAATTTTCCCTTTTAAAGGGGGGAGTCCCTTTTCTATCTTGGGGTGGGATTTGTG 2040
D 1981 AGAATTTTCCCTTTTAAAGGGGGGAGTCCCTTTTCTATCTTGGGGTGGGATTTGTG 2040
QY 2041 GCGCCCAACGAGGAGCTTTATTTATTTATGATGAAATCTGGCGATTTTGGCCCTA 2100
D 2041 GCGCCCAACGAGGAGCTTTATTTATTTATGATGAAATCTGGCGATTTTGGCCCTA 2100
QY 2101 CTTATTTACAGCGCCCGCAAAATTTAAACGAAGAACTTAAGTATGGCGGT 2160
D 2101 CTTATTTACAGCGCCCGCAAAATTTAAACGAAGAACTTAAGTATGGCGGT 2160
QY 2161 ACAATTTACCAACACCTGTGAGAGGAGTATCCGCCCTCAATACACTTTCTGTTC 2220
D 2161 ACAATTTACCAACACCTGTGAGAGGAGTATCCGCCCTCAATACACTTTCTGTTC 2220
QY 2221 CCGCAAAATGGAATAATGATCCGAGAGGCCCATCTTTTGGGGGAAGTATCACTA 2280
D 2221 CCGCAAAATGGAATAATGATCCGAGAGGCCCATCTTTTGGGGGAAGTATCACTA 2280
QY 2281 TTACCTTTACCAAGACTATCCAGATGGAATGTAAGGATGGGGCATGCAAGCTC 2340
D 2281 TTACCTTTACCAAGACTATCCAGATGGAATGTAAGGATGGGGCATGCAAGCTC 2340
QY 2341 CGAAGATTTTATGATGAGAGGAGGATTCATCCGGAAGTATACCAATAAATA 2400
D 2341 CGAAGATTTTATGATGAGAGGAGGATTCATCCGGAAGTATACCAATAAATA 2400
QY 2401 TGGTGATC 2408
D 2401 TGGTGATC 2408

RESULT 2

US-09-986-682B-4
Sequence 4, Application US/09986682B
Patent No. US20020115182A1

GENERAL INFORMATION:

APPLICANT: KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAKAKU
KENKYUJO

TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE
ACTIVITY

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:

ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20004

COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/986,682B
FILING DATE: 06-No. US20020115182A1-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 170,630/1996
FILING DATE: 10-JUN-1996

ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.

REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 1365 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-986-682B-4

Query Match 56.7%; Score 1365; DB 10; Length 1365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 ATGAACAGCGGGAGCTACAGAGAACTATGTTTGGCCATATTACAGCGCTGACATG 516
D 1 ATGAACAGCGGGAGCTACAGAGAACTATGTTTGGCCATATTACAGCGCTGACATG 60
QY 517 CTAATAATTCAGAGCAACAACAGTCCCAATTTAAAGAGCCCTCAATTCATCA 576
D 61 CTAATAATTCAGAGCAACAACAGTCCCAATTTAAAGAGCCCTCAATTCATCA 120
QY 577 GCAATCAAAAACATTGATTCGGCAAAAGGATGATAGTACAGCACTTAATGATTTA 636
D 121 GCAATCAAAAACATTGATTCGGCAAAAGGATGATAGTACAGCACTTAATGATTTA 180
QY 637 GATGATGAGGATAGCTGCGCACTGCAAAACGCTGATGCTGCGCAATTTATCAGGA 696
D 181 GATGATGAGGATAGCTGCGCACTGCAAAACGCTGATGCTGCGCAATTTATCAGGA 240
QY 697 TATCAGATCGTCTCGCTTTAGCAGGTGACCCCAAAAACAGTGAATGATCTCCACTTCA 756
D 241 TATCAGATCGTCTCGCTTTAGCAGGTGACCCCAAAAACAGTGAATGATCTCCACTTCA 300
QY 757 TTATTTATCAAAAAGTCTGATATACATGATGACAGCTGGAATAATGCTGAAGAGTA 816
D 301 TTATTTATCAAAAAGTCTGATATACATGATGACAGCTGGAATAATGCTGAAGAGTA 360
QY 817 TTTGAGATATGATTAATTTGTTCCAAATATGATCCGTAATTAATATCAACACAGAG 876
D 361 TTTGAGATATGATTAATTTGTTCCAAATATGATCCGTAATTAATATCAACACAGAG 420
QY 877 TGGTCAGGTTCTGCTACTTTAACCAAGATGGCAAGCGCTTATCTATACAGATTAC 936
D 421 TGGTCAGGTTCTGCTACTTTAACCAAGATGGCAAGCGCTTATCTATACAGATTAC 480
QY 937 TCAGTAATCTCGAAGATGATGTAACCGGCTGCTGTAACCAATATTCACATGCTCAA 996
D 481 TCAGTAATCTCGAAGATGATGTAACCGGCTGCTGTAACCAATATTCACATGCTCAA 540
QY 997 GTAACCTATCCAGCGGAGTACAGCTACCTTAAAGTGAAGATGATGATCA 1056
D 541 GTAACCTATCCAGCGGAGTACAGCTACCTTAAAGTGAAGATGATGATCA 600
QY 1057 TCTGCTTTGATGGCGGAGAGGTATGATCAAAATATTCAGCAATTTATGATGAA 1116
D 601 TCTGCTTTGATGGCGGAGAGGTATGATCAAAATATTCAGCAATTTATGATGAA 660
QY 1117 GCGAAGTGAATTTAGGATTAACCACTTTAAGAGCCCTCACTATGTTGAAGATTA 1176
D 661 GCGAAGTGAATTTAGGATTAACCACTTTAAGAGCCCTCACTATGTTGAAGATTA 720
QY 1177 GCGCATAAATATCTTGTCTTTGAAGCGAATACGTAACACAGATGTTATCAAGCGAT 1236
D 721 GCGCATAAATATCTTGTCTTTGAAGCGAATACGTAACACAGATGTTATCAAGCGAT 780
QY 1237 CAGCTTTCAATTAATAAGCTTACTATGCGGAGTACAGCTCTTCCCAATGAATAA 1296
D 781 CAGCTTTCAATTAATAAGCTTACTATGCGGAGTACAGCTCTTCCCAATGAATAA 840
QY 1297 AATAAAGCTGTTCAAGTCTTAATAAACAATTTCTTTAGGAAATGATGATGATG 1356
D 841 AATAAAGCTGTTCAAGTCTTAATAAACAATTTCTTTAGGAAATGATGATGATG 900
QY 1357 ATTGTTGAATTTGCGGATGACTATACAGTGAAGTGTATGAACAATTTAGTCATCA 1416
D 901 ATTGTTGAATTTGCGGATGACTATACAGTGAAGTGTATGAACAATTTAGTCATCA 960

QY	601	AAAGGATGATGAAGCAGGCAACTTATTAATTTAGATGATGGGTACTGGCCAGT	660
Db	232	AAAGGATTAACAAAATAATGGAGAGCTGATTCATTTAGACGTATGGGCAAGCTGGCCGTG	291
QY	661	CAAAACGCTGATGCTACTGCGGCAAAATATTCATGATATCCATCGTCCGCTTAGCA	720
Db	292	CAAAATGCGGAGCGGGACGGTTGCTCATATACCAGGCTACATCTGTTTTGCGGCTGGCG	351
QY	721	GGTGACCCAAAAAACAGTGTATATCTCACTTATTATTCTATCAAAAAGTCGGTGAT	780
Db	352	GGGATCCCAAAAGACGTGATGACACATCCATCTATTGTGTCATCAAAAAGAAAGCGAA	411
QY	781	ACATCATTTAGACAGCTGGAAAAATCTGTGAAGAGATTTGGAAGTATGATTAATTTGTT	840
Db	412	ACTTCTATGACACGCTGGAAAAAGCCGGCAGAGTGTTTAAGACACGCAACAAATTTGTT	471
QY	841	CCAAATGATCCGATCTTAAATATCAAAACACAGAGAGTGTCAGGTTCTGCTACTTTAAC	900
Db	472	CCAGACGATCCGATACCTCAAAATCAAAACACAGGAATGGTCAGTCTGCCACGCTGACA	531
QY	901	AAAGATGCGCAAGTCGTTTATTCTATACAGATTACTCAGTAATCTGAAGATGGTGA	960
Db	532	AAAGACGGAAGAAGTCCGACTGTTTACACACACTTTTTCGCGACGCAATAA-----	581
QY	961	ACCGGTGCTGATGATCCAAATCATTTCACTGCTCAAGAAACCTATCCGACGGCGGATGCA	102
		582	-----CGGACAGCAGCGCTGACCAACACTGAGTCAATTTCTCTACGCCGAGTTTCG
		633	

0-2644

QY 1081 ACAGCTTATCAAAATTATTCAGCAATTATTCGATGTGAAGGCAAGTGATTTCCAGGTGATTAAC 114
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 694 ACGGTATACCAAAACGCTTCAGCAATTCATTCACGAGAAGAACTTAGACTTCGGCGACAAC 753
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1141 CATACCTTTAAGAGACCCCTCATTGTGTTGAAGATTAAGGCCATAAATATCTTGTCTTTGAA 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 754 CATACGATGAGAGACCCGCATTATGTGGAAGACCGCGCCATAAATATCTGATTTTGA 813
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1201 GCGAATACCTGGAACCAACAGATGCTTATCAAGCGCATGAGCTTTCATATATAAGCTTC 126
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 814 GCCAATATCGGGAACAAAAACGGCTACCAAGAGAAGCTCCATTCCAACAGACCTTAC 873
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1261 TATGCGGAAGAGACGCTCTTCTTCGAGATGAAAAAATAACTGCTTCAAGCTCTAAA 132
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 874 TAGGGGGGACAGAAAGTTCTTTAAAGAAGAAAGCAAGCAAGCTGCAAGGTGGGAAC 933
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1321 AAACAATATGCTTCTTTAGCGATGCTGCTATTAGGCATTTGTGAATTTGGCCGATGACTAT 138
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 934 AAAAACAACGCTTCGCTGGCTAACGGGCGCTCTCGAATCATCGAATTAATAATACGATTAAT 993

Db

Db	1054	CGGGCCACCTCTTCAAAATGATGAAATGGTATCTGTTCCACAGATTCAAGAGATCA	111
QY	1501	AAAATGACGAGTAGTGAAATTAAACGACAAGAATGTTTATATGCTAGGAGCCGAGGCGAC	156
Db	1114	AAATGACAAATTGACGGCATGCGTTCANAAGCAATTATATGCTAGCTATGATCAAGT	117
QY	1561	TCCTTAAATGCGCCACACACACCCGATTAATGATAAACTGACTGTTATTGAACATGATCTT	162
Db	1174	TCATTAAACCGGACCAATTCAAGCTTTAAACAACACCGACTGTTTGGCATATGAGACAG	123
QY	1621	GACCCGCTGCATCTCACACACACTTACATCTGATGCGGTATCCCGCACCCCTGAAGGTAAAT	168
Db	1234	GATTACAAATGACATACAGCTTACTTATTACACACTTTGCGGTACCGCAGAAAAAAGCGGAC	129

OY		898	ACCAAGTGGCCAAAGGCCGTTTATTCATATACATTAACAGGAATCCTGAAGATGCT	957
Dd		2016	ACAATCGCGAATAAATCCGTTTATTCTACACTATTTCTCGGTAAACATTA-----	2068
OY		958	GGAACCGGTGCTGGTAAACCAATATATTTCAACTGCTCAAGTAAACTTATCCAGCCGGAT	1017
Dd		2069	-----CGGCAAAACAACCTGCACACTCACACAGTTAACGTATGCATCGCATCGAC	2117
OY		1018	GCAGCTACCTTAAGTCGATGAGATATGATATATAATCTGCTTGATGGCGGAGAC	1077
Dd		2118	AGCTCT---TTGAACATCAACGGTGTACAGGATTAATTAATCACTTTGA---CGGTGAC	2171
OY		1078	GGTACAGTTTATCAAATAATTTACAGCAATTTATCGATGAAGCCAAGTGATTTACGTCAT	1137
Dd		2172	GGAAAAACGPTATCAAAATGTCACACAGTTCATCGATGAAGCCAACCTACAGTCAAGCGAC	2231
OY		1138	AACCATCTTAAAGGACCCTCAGTATBTTAAGCTAAAGGCCATTAATATCTTGCTTT	1197
Dd		2232	AACCATGCGTGNAGATCTCTCACTACGTAGAAGATTAAGGCCCAAAATACTTGATATT	2291
OY		1198	GAAGCGAATCTGGAACAACAGATGTTATCAAGCGCATAGCTCTTTCATATPAAGCT	1257
Dd		2292	GAAGCAAAACCTGGAACACTGAAGATGGCTACCAAGCGCAAGATCTTTATTTAAACAAGCA	2351
OY		1258	TACTATGCGGGAAGTAGACGCTCTCTCCAGATGSAAAAATTAACGTCTTCAAGAAGTCT	1317
Dd		2352	TACTATGGCAAAACACATCATTTCTTCGCTCAAGAAATCAAAAACCTTCTCAAAACGAT	2411
OY		1318	AAAAAACAAATTGCTTTCTTTAGCGAATGTCATTAAGGCATTTGTGAATTGGCCGATGAC	1377
Dd		2412	AAAAAACCGACGGCTGAGTATAGCAAAACGGCCCTCGGTATGATTTAGGACTAAACGATGAT	2471

```

; TITLE OF INVENTION: COMPOSITIONS FOR USE IN PRACTICING THE SAME
; FILE REFERENCE: CLON-069
; CURRENT APPLICATION NUMBER: US/10/055,794
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/263,358
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Vector
US-10-055-794-3

Query Match      25.3% Score 608.8; DB 9: Length 6525;
Best Local Similarity 66.4%; Pred. No. 1,5e-152;
Matches 994; Conservative 0; Mismatches 447; Indels 57; Gaps 6;

OY 358 ACGATGACCTTCAAAAGATTGGCGAAAGACGCGTACCTTCAGACCTGCTATA 417
DB 3093 ACGATGACCTTCAAAAGATTGGCGAAAGACGCGTACCTTCAGACCTGCTATA 417
OY 418 TTAGTAGAGCGGACGACGCCATATTTTGGCCAGCAAGTGAACAGCGGGGACTACAG 477
DB 3153 CTGGCAGAGAGG---CGCAACTCAGCGCTTGGCAAGAACGACCAAAAGCCATATTAAG 3209
OY 478 GAAGCTATAGTTGGTGGCCATATTACAGCGCTGACATGCTAAATTCGAGCAACAA 537
DB 3210 GAAGCTATAGTTGGTGGCCATATTACAGCGCTGACATGCTAAATTCGAGCAACAA 537
OY 538 AACAGTCCCTCAATTTAAAGTCCCTCAATTTCAATGATGATGATGATGATGATG 597
DB 3270 AAAAATGAAATAATCAAGTTCCTGATGCTGATGCTGATGCTGATGCTGATGCT 3329
OY 598 GCAAAAGGCTATGATTAAGTCAAGCACTTAATAGATTAGATGATGATGATGATG 657
DB 3330 GCAAAAGG---CTTGAGCTGTTGGGACAGCTGGGCA 3362
OY 658 CTGCAAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
DB 3363 TTACAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
OY 718 GCAGGTGACCCAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATG 777
DB 3423 GCCGAGATGCTAAAGATGCGGATGACATGATGATGATGATGATGATGATGATG 3482
OY 778 GATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
DB 3483 GAAACTTCTATTGACAGCTGGGAAAGAGCTGGCGCTTTAAAGACGCGCAAAATTC 3542
OY 838 GTTCCAAATGATCCGTATCTTAATATCAACACAGAGTGGTGGTCTGCTCTTTA 897
DB 3543 GATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3602
OY 898 ACCAAAGTGGCGAGTCCGTTTATTTATATACAGATTAATGATGATGATGATGATG 957
DB 3603 ACATCTGACGGAAGAAATCCGTTTATTTACACTGATTTCTCGGTAACATTA----- 3655
OY 958 GGAACCGGTGCTGGTACCAATCTTTTCACTGCTCAAGTAACTTATCCAGCGGAT 1017
DB 3656 -----CGGCAACCAAACTGACACACTGACCAAGTAACTTATCCAGCATCAGAC 3704
OY 1018 GCAGCTACACTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1077
DB 3705 AGCTT-----TTGACATTAACGCTGTAGAGATTAATATCAATCTTTGA---CGGTGAC 3758
OY 1078 GGTACAGTTTCAAAATATTCAGCAATTTATGATGATGATGATGATGATGATGATG 1137
DB 3759 GGAAGAAAGTATCAAAATATGATGATGATGATGATGATGATGATGATGATGATG 3818
OY 1138 AACCACTATTAGAGACCCCTCCTATGTTGAAGTAAGGCCATTAATATCTTCTTT 1197

```

```

DB 3819 AACCATACCTGAGAGATCTCCTACTACTAGATTAAGGCCAACAAATACTGTATT 3878
OY 1198 GAAGCAATGATGAGCAACAGATGATGATGATGATGATGATGATGATGATGATGATG 1257
DB 3879 GAGCAAACTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 3938
OY 1258 TACTATGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1317
DB 3939 TACTATGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3998
OY 1318 AAAAACAATGCTCTTTTACGCAATGCTCATTAGCATGTTGATGATGATGATGATG 1377
DB 3999 AAAAACAATGCTCTTTTACGCAATGCTCATTAGCATGTTGATGATGATGATGATG 4058
OY 1378 TATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1437
DB 4059 TACACACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4118
OY 1438 GAACCGCCCAATATATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1497
DB 4119 GAACCGCCCAATATATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 4178
OY 1498 TCACAAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557
DB 4179 TCACAAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4238
OY 1558 GACCTCTTAATGAGCCACACACCGATTAATGAATGAATGAATGAATGAATGAATGAAT 1617
DB 4239 AATCTTTAAGTGGCCATACAGCGCTGACCAAAATGCTGCTGCTGCTGCTGCTGCTG 4298
OY 1618 CTGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1677
DB 4299 CTGATGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4358
OY 1678 AATAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1737
DB 4359 AACAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4418
OY 1738 CACCTGCGGACCACTGCGGATTAATGAATGAATGAATGAATGAATGAATGAATGAAT 1797
DB 4419 ACGTTTGGCCTAGCTCTCTGCTGACATCAAGGCAAGAAACATCTGTTGTCAAGAC 4478
OY 1798 AGTTCC---GGACAAGCAATTCCTCATGAGGATTAATCCCAATTAAGAAATGTC 1852
DB 4479 AGCATCTTGAACAAGCAATTAACAGTTAAACAATTAAGCAAGAAAGAAATGCC 4536

RESULT 6
US-09-849-866-1/c
; Sequence 1, Application US/09849866
; Patent No. US200206833A1
; GENERAL INFORMATION:
; APPLICANT: Ilya Chumakov
; Hitoaki Tanaka
; TITLE OF INVENTION: High Throughput DNA Sequencing Vector
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobel, Martens, Olson & Bear, LLP
; STREET: 550 West C Street, Suite 1200
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/849,866
; FILING DATE: 04-May-2001
; ATTORNEY/AGENT INFORMATION:

```

NAME: Daniel Hart
 REGISTRATION NUMBER: 40,637
 REFERENCE/DOCKET NUMBER: GENSET.15CDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10317 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: DOUBLE
 TOPOLOGY: CIRCULAR
 MOLECULE TYPE: synthetic DNA
 ORIGINAL SOURCE:
 ORGANISM: Cloning vector pGendEL
 FEATURE:
 NAME/KEY: pGendEL
 LOCATION: 1..10317
 FEATURE:
 NAME/KEY: Homology with X06404 compl (411..1668)
 LOCATION: 9..1266
 IDENTIFICATION METHOD: blastn against X06404
 FEATURE:
 NAME/KEY: Kanamycin resistance gene CDS
 LOCATION: 142..957
 IDENTIFICATION METHOD: By homology to X06404
 FEATURE:
 NAME/KEY: Tn1000's right end
 LOCATION: complement 1332..1371
 IDENTIFICATION METHOD: blastn against X60200)
 FEATURE:
 NAME/KEY: Homology with U46017 (1-472)
 LOCATION: 1423..1894
 IDENTIFICATION METHOD: blastn against U46017
 FEATURE:
 NAME/KEY: single stranded DNA replication origin
 LOCATION: 1423..1894
 IDENTIFICATION METHOD: By homology to U46017
 OTHER INFORMATION: mutation T -> C 1658
 FEATURE:
 NAME/KEY: Homology with U51113 (2382..6997)
 LOCATION: 1896..6544
 IDENTIFICATION METHOD: blastn against U51113
 FEATURE:
 NAME/KEY: OriS
 LOCATION: 1972..2188
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: repELR
 LOCATION: 2897..2918
 OTHER INFORMATION: Described in seqid 16
 FEATURE:
 NAME/KEY: RepE
 LOCATION: 2903..3034
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: T3
 LOCATION: 3043..3059
 OTHER INFORMATION: Described in seqid 17
 FEATURE:
 NAME/KEY: LRT3RA
 LOCATION: complement 3045..3069
 OTHER INFORMATION: Described in seqid 15
 FEATURE:
 NAME/KEY: Incc
 LOCATION: 3070..3320
 IDENTIFICATION METHOD: By homology to U51113
 OTHER INFORMATION: Insertion 33 bases 3038..3071
 FEATURE:
 NAME/KEY: Para
 LOCATION: 3655..4821
 IDENTIFICATION METHOD: By homology to U51113
 OTHER INFORMATION: mutation G -> A 3878

FEATURE:
 NAME/KEY: ParaB
 LOCATION: 4821..5792
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: ParC
 LOCATION: 5865..6382
 IDENTIFICATION METHOD: By homology to U51113
 FEATURE:
 NAME/KEY: Homology with J01688 (complement 175..819)
 LOCATION: 6574..7218
 IDENTIFICATION METHOD: blastn against J01688
 OTHER INFORMATION: mutation A -> G 7096
 FEATURE:
 NAME/KEY: CDS streptomycin sensitivity gene
 LOCATION: complement 6716..7090
 IDENTIFICATION METHOD: By homology to J01688
 OTHER INFORMATION: mutation A -> G 6728
 mutation G -> C 6821
 mutation C -> T 6866
 mutation T -> C 7013
 mutation T -> A 7058
 FEATURE:
 NAME/KEY: tpsLR
 LOCATION: 7155..7174
 OTHER INFORMATION: Described in seqid 12
 FEATURE:
 NAME/KEY: SP6
 LOCATION: 7230..7248
 OTHER INFORMATION: Described in seqid 13
 FEATURE:
 NAME/KEY: Tn1000's left end
 LOCATION: 7252..7291
 IDENTIFICATION METHOD: blast (X60200)
 FEATURE:
 NAME/KEY: Homology with X02730 (complement 37..1959)
 LOCATION: 7305..9227
 IDENTIFICATION METHOD: blastn against X02730
 FEATURE:
 NAME/KEY: CDS levansucrase gene
 LOCATION: complement 7379..8800
 IDENTIFICATION METHOD: By homology to X02730
 OTHER INFORMATION: mutation T -> C 7466
 mutation A -> G 7739
 mutation T -> C (Asn -> Asp) 8347
 mutation T -> C 8600
 mutation G -> A (Ala -> Val) 8772
 FEATURE:
 NAME/KEY: SLR3
 LOCATION: 8711..8731
 OTHER INFORMATION: Described in seqid 14
 FEATURE:
 NAME/KEY: Homology with J01636 (complement 1158..1465)
 LOCATION: 9298..9623
 IDENTIFICATION METHOD: blastn against J01636
 FEATURE:
 NAME/KEY: CDS alpha peptide beta-galactosidase
 LOCATION: complement 9276..9497
 IDENTIFICATION METHOD: By homology to J01636
 FEATURE:
 NAME/KEY: primer HE1
 LOCATION: complement 9465..9479
 FEATURE:
 NAME/KEY: primer HE2
 LOCATION: 9461..9475
 FEATURE:
 NAME/KEY: primer lacRS2AVT
 LOCATION: complement 9603..9630
 FEATURE:
 NAME/KEY: primer lacE2mlu
 LOCATION: 9289..9314
 FEATURE:
 NAME/KEY: Homology with M77789 (1889..2576)


```

1 LOCATION: 9629..10315
2 IDENTIFICATION METHOD: blastn against M77789
3
4 FEATURE:
5 NAME/KEY: high copy-number double-stranded DNA replication origin
6 LOCATION: complement 9629..10315
7 IDENTIFICATION METHOD: By homology to M77789
8 OTHER INFORMATION: mutation C -> T 9803
9 site Pml1 10038 - 10034
10 CLONING SITES 10031 - 10041
11
12 FEATURE:
13 NAME/KEY: oriLRd
14 LOCATION: 9856..9881
15 OTHER INFORMATION: Described in seqid 8
16
17 FEATURE:
18 NAME/KEY: OSI
19 LOCATION: 10009..10026
20 OTHER INFORMATION: Described in seqid 10
21
22 FEATURE:
23 NAME/KEY: ORI
24 LOCATION: complement 10046..10062
25 OTHER INFORMATION: Described in seqid 11
26
27 FEATURE:
28 NAME/KEY: oriLRt
29 LOCATION: complement 10182..10202
30 OTHER INFORMATION: Described in seqid 9
31
32 SEQUENCE DESCRIPTION: SEQ ID num: 1:
33
34 OS-09-849-866-1

```

Query Match	25.28:	Score	607.2:	DB:10:	Length	10317:			
Best Local Similarity	66.38:	Pred. No.	5.5e-152:						
Matches	993:	Conservative	0:	Mismatches	448:	Indels	57:	Gaps	6:
OY	358	ACGATGACCTTCAAAAGATTGGCGGAAAAAAGACAGCTGCCGTACCTTCACAGSACTGCTATA	417						
Db	8803	ACGATGAACATCAAAAAGTTTGGCAAAACAGTAAGATTAACCTTTACTACCGGACAG	8744						
OY	418	TTAGTAGACGCGAGSAGCCGATATTTTGGCAGCAATGAACAGGGGACTCAAG	477						
Db	8743	CTGGCAGGAGG--CGCACTCAAGCGTTTGGCAAAAAGAAACGAAACCAAAAGCCATTAAG	8687						
OY	478	GAAACATGATGTTTGGCCATATTACACGCGCTGACATGCTAAAAATTCAGGACAAACA	537						
Db	8686	GAAACATACGCGCATTTCCCATATTACGCGCATGATATGTGCAATTCCTGACAGCA	8627						
OY	538	AACAGTCCCTAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAACATTGATCG	597						
Db	8626	AAAAATAAAAATATCAAGTTCCTGAGTTGATTCGTCACAAATTTAAATATCTCTCT	8567						
OY	598	GCAAAAGSGATGATTAAGTCAGCACTTAATGATTTAGATTAAGGATAGCTGCGCA	657						
Db	8366	GCAAAAGS-----CCTGAGCTTTGGGACACTGCGCA	8534						
OY	658	CTGCAAAACGCTGATGCTACTGCGGCAAAATATCATGATGATACATCGTCCGGTTA	717						
Db	8533	TTACAAAACGTGACGCGACTGTCGCAAACTATACAGGGTACACATCGTCTTGATTA	8474						
OY	718	GCAGGTACCCAAAACAAAGTATGATATCTGCACCTTCATTATTCTATCAAAAAGTCGT	777						
Db	8473	GCCGGAATCTTAAAAATGGGAGTACACATCATGATTTACATGTTCTATCAAAAAGTCGC	8414						
OY	778	GATCATTCGATGACAGCTGGAATAATGCTGGAAGATATTGAAGATATGATTAATTT	837						
Db	8413	GAACTTCATATGACAGCTGGAATAATGCTGGAAGATATTGAAGATATGATTAATTT	8354						
OY	838	GTTCCAAATGATCGTATCTTAATATCAAAACAGAGATGTCAGATTCCTGCTACTT	897						
Db	8353	GATCAGATGATTCATCTATAAAGACCAACACAAAGATGTCAGATTCAGCCATTT	8294						
OY	898	ACCAAAATGCGCCAGTCCTTTATTCTATACAGATTACTCAGATAATCCTGAAGATGGT	957						
Db	8293	ACATCTGACGGAATAATCGTTTATTTCTACATGATTTCTCGGTAAACATTA-----	8241						

QY	958	GGAAACGGTGGCTGATACCAACCAATCACTTTCACACTGGCTCAAGTAACCTATTCACCGCGGAT	107
Db	8240	-----CGGAAACCAACACTGACAACTGGACAGTTAAAGCTATACACATCAGAC	819
QY	1018	GCAGCTACACTTAAAGTCGATGGAGTAATCTCATCTTAATCTGCTTTGATGGCGGAGAC	1077
Db	8191	AGCTCT---TTGAAATCATCAGCGGTGTAGAGGATTTAAATCATCACTTTTGA---CGGTGAC	8138
QY	1078	GGTACAGTTTATCAAAATATTTCAGCAATTTATCCATGAAAGGCAAGTGGATTTTCAGTGT	1137
Db	8137	GGAAAAACGTTACAAATGTTACAGGAGTTCACTCGATGGAAGGCAACTACAGCTCAGCGAC	8078
QY	1138	AACCTACTTTAAGAGACCTCTCATGTGTTGAAGATTAAGGGCCATAATATCTTCTTT	1197
Db	8077	AACCTACCGCTCAGAGATCTCTCAGTACAGAAAGTAAAGGCCAACAAATACTTACTATTT	8011
QY	1198	GAAAGCAATCTGGAACAACAGATGTTATCAAGCGCATAGCTTTTCAATATTAAGCT	1257
Db	8017	GAAAGCAACACTGGAAGTACAGATGCTTACCAAGCGCAAGATCTTTTAAACAAACA	7958
QY	1258	TACTATGGCGGAAGTACGCTCTCTCCAGATGAAAAAATAAATCACTGCTCAAGTCT	1311
Db	7957	TACTATGGCAAAAGCACATCATTTCTTCGCTCAAGAAAGTCAAAAACCTCTCGAAGCGAT	7899
QY	1318	AAAAACAATATGCTCTTTAGCCGAATGGTGCAATTAGGCATTTGGAATTGGCGGATGAC	1377
Db	7897	AAAAAACGACGGCTGAGTTACCAAAACGGCGCTCTCGGATGATTTAGACTTAAACGATGAT	7838
QY	1378	TATACAGTGAAGAAGTATTGAACACATTAGTCGCATCAACACACGTAGCAGATACAGTC	1433
Db	7837	TACACACGTAAAAAAGTATGTAACACCGCTGATTGATCTTAACACACGTAAACAGATATAAT	7777
QY	1438	GAAACCGCCAAATATATTTAAATGATATATAATGATATCTATTCACGGATTCAGAGAGA	1497
Db	7777	GAAACCGCCAAACGCTTTTAAATGAAAGCAAAATGATATCTGACAGTACCTCCGCGGA	7712
QY	1498	TCCAAATATGACGATGATGGAATTACGACAAGATGTTTATATGCTTAGGCGCCGGAAGC	1557
Db	7717	TCAAAAATGACCATTTGACCGCATTCGCTAACGATATTTACATCTCTGGTATATTTCT	7658
QY	1558	GACTCCTTAAATGGGCCACACAAACCGGATTAATGAACTAGGACTGTATTGAACTGAT	1611
Db	7657	AATCTTTTAACTGCGCCCATACAGCGCTGAACAAAACTGGCCTTGCTGTATAAAAATGAT	7599
QY	1618	CTTGAACCTGCTGATTCACACACACTTACTCTCAATGGGATATCCGCAACCTGAAGGT	1677
Db	7597	CTTGAATCCCAAGCATGATTAACCTTTACTTACTACACTTCGCTGATCTCTCAACCGGAAGA	7538
QY	1678	AATATGTGTACTCTGACAAAGTATATATGAGAAATAGAGGCTCTCTACGAAACATACGCT	1733
Db	7537	AACAAATGTGTGATTAACAAGCTATATGACAAACABAAGGATTTCTACGCAACAAACAATCA	7477
QY	1738	CACCTGCGGAGACAACTTGGGGTATATATTAAGAGGTCTGACACATCTGAGAGAGAAAT	1797
Db	7477	ACGTTTGGGCGGAGCTTCCTGCTGAAATCAAAAGGCAAGAAACAATCTGTCTCAAAAGAC	7411
QY	1798	AGTTCC---GGAAGAAGCAATTTCCCATATGCGATATATCTCCCAATAAAGCAATGTC	1852
Db	7417	AGCATCTCTTGAACAAGGACAAATTAACAGTTAAACAATATAAAACGCAAAAGAAATGCC	7360

```

RESULT 7
US-09-993-292A-19
; Sequence 19, Application US/09993292A
; Patent No. US20020146430A1
; GENERAL INFORMATION:
; APPLICANT: James E. Galen
; APPLICANT: University of Maryland
; TITLE OF INVENTION: USE OF C1Y A HEMOLYSIN FOR EXCRETION OF
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: USFMD.007A
; CURRENT APPLICATION NUMBER: US/09/993.292A
; CURRENT FILING DATE: 2001-11-23

```


575 CAGCAATCAAAAACATGTTGTCGCGCAAAAGGATGATAGTACGAGCAACTAATAGATT 634
1555 CCACAAATTAATAATATCTCTCTGCAAAAG-----CC 1587
635 TAGATGATGGAGATGCTGGCCACTGCAAAAGCTGATGTAAGTCCGCCAAATTAATCAG 694
1588 TGGAGTTGGGAGCTGGCCATTAACAAAGCTGAGGCGCACTGTCCAAACTATCAG 1647
695 GATATCATATGCTCCGCTTACAGAGTGAACCCAAAAACAGTATGATGATCTCAGTTC 754
1648 GCTACCATGATGCTTGTGATGAGGAGATCTTAATAAATGCGATGACATGATGATT 1707
755 ATTTATTTATCAAAAAGCTGGTGTATCATGATGACAGCTGGAATAATGCTGGAGAG 814
1708 ACATGTTCTATCAAAAAGCTGGCGCAAACTTCATTGACAGCTGGAAAAAGCTGGCCGC 1767
815 TATTTGAAGATATGATTAATTTGTTCCAAATGATCCGTATCTTAATATCAACACAG 874
1768 TCTTTAAAGACAGCGCAAAATTCGATGCAAAATGATCTTATCTTAAGACCAAG 1827
875 AGTGTGAGTTCTGCTACTTAAACCAAGATGCGCAAGTCCGTTATTTATAGAGATT 934
1828 AATGTCAGGTTACCCACATTACATCTGACGGAATAATCCGTTATTTACACTGATT 1887
935 ACTCAGTAACTCCTGAAGATGCTGGAACCGGTGCTGTAACCAATTCATTTCACTGCTC 994
1888 TCTCGGTAACATTA-----CGGCAACAAACACAGCAACTGACAC 1929
995 AAGTAACTTATCCACGCGGATGACCTACACTTAAAGTCGATGAGTATGATCAATA 1054
1930 AAGTTAAGTATACACATCAGACACCTCT---TTGAACATCAACGCTGAGAGATTATA 1986
1055 AATGTCGTTGATGCGGAGCGGTACAGTTTATCAAAATTTAGCAATTTATCAGATT 1114
1987 AATCATCTTTTA---CGGTACGGAATAAAGCTATCAAAATGTACAGAGTTTCATGATG 2043
1115 AAGCAAGTGTATTCAGGTGATTAACATCTTAAAGAGCCCTGACTATGTTGAAGATA 1174
2044 AAGGCACTACAGCTCAGCGGCAACCAATACCTGAGAGATCTCTACGTAAGAGATA 2103
1175 AGGCAATTAATATCTGCTTTGTAAGCAATACGTAACCAACATGTTATCAAGCGC 1234
2104 AAGGCAACAAATCTAGTATTTGAAGCAAAACGTAAGTGAAGTCTTACCAAGCGC 2163
1235 ATCAGTCTTTCAATATAAGCTTACTATGCGGAGTACGCTCTTCTCCAGATGATA 1294
2164 AAGATCTTTTATTAACAAGATATCTATGCAAAAGCAGCATCTTCTCCGTCAGAGATA 2223
1295 AAAATTAACGCTTCAAGTCTTAATAAACAATGCTTCTTACGGAATGTCGATTAG 1354
2224 GTCAAAAACCTTCTGCAAGCGATTAATAAAGCAGCGCTGAGTTACCAAGCGCTCAG 2283
1355 GCATGTTGATTTGGCCGATGCTATACAGTGAAGAGTGTATGAACCATTAAGTCGAT 1414
2284 GTATATATGAGTAAACGATGATTTACACAGTGAAGAGTGTGAAGCGCTGATGATGAT 2343
1415 CAACACAGTACAGATGAAGTGAAGCGGCAATATATTTAAATGAATATAATAGT 1474
2344 CTAACACAGTACAGATGAATGAAGCGGCAAGCTCTTTAAATGAAGCGCAATAGT 2403
1475 ATCTATGACGATTCAGAGAGATCAAAATGACAGTGAAGTGAATTAAGCAAGATG 1534
2404 ACCTGTTACTGACTGCTCCGCGGATCAAAAATGACAGTGAAGGCAATTAAGTGAAGATA 2463
1535 TTTATATGCTAGGCGCGGAGCGCTCTTAATGCGCCACACAAACCGATTAATGATA 1594
2464 TTTATATGCTGCTGTTCTTCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2523
1595 CTGAGCTTGTATGTAACATGATCTGACCCGCTGATGCTGACACAGCTTCTCTCTCT 1654
2524 CTGAGCTTGTATTAATAATGATCTGATCTTAACAGATGATGATCTTACTTACTGACACT 2583

QY 1655 GCGGATCCCGCACCTGGAAGGATATATGTTGTTACTACAGATTAATATGACGAATAGAG 1714
DB 2584 TCGGCTGATCTCAAGGCAAGGAAAGCAATCTGCGATTACAGCTATATGACAAACAGAG 2643
QY 1715 GCTTCTATCCAGAACATCATCTTCACTGCGGAGCAAGCTTGGGTTAATTTAAAGGCT 1774
DB 2644 GATTTCTAGCGAGCAAAACAAATCAACGTTTGGCCCAAGCTTCTGCTAATCAAAAGCA 2703
QY 1775 CTGACACATCTGAGGAGGAATAATAG 1799
DB 2704 AGAAACATCTGTTGTCAAGAGAGAG 2728

RESULT 9

US-09-864-761-21746
Sequence 21746, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-x-1
CURRENT FILING DATE: US/09/864, 761
PRIORITY FILING DATE: 2001-05-23
PRIORITY FILING DATE: US 60/180, 312
PRIORITY FILING DATE: 2000-02-04
PRIORITY FILING DATE: US 60/207, 456
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: US 09/632, 366
PRIORITY FILING DATE: 2000-08-03
PRIORITY FILING DATE: US 60/263, 6
PRIORITY FILING DATE: 2000-10-04
PRIORITY FILING DATE: US 60/236, 359
PRIORITY FILING DATE: 2000-09-27
PRIORITY FILING DATE: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: US 60/234, 687
PRIORITY FILING DATE: 2000-09-21
PRIORITY FILING DATE: US 09/608, 408
PRIORITY FILING DATE: 2000-06-30
PRIORITY FILING DATE: US 09/774, 203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21746
LENGTH: 368
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035452.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.6

```
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
OTHER INFORMATION: NT HIT: X52988.1, EVALUE 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: P21130, EVALUE 1.00e-64
OTHER INFORMATION: EST_HUMAN HIT: BF219305.1, EVALUE 4.00e-07
NAME/KEY: unsure
LOCATION: 332
NAME/KEY: unsure
LOCATION: 335
NAME/KEY: unsure
LOCATION: 356
NAME/KEY: unsure
LOCATION: 357
NAME/KEY: unsure
LOCATION: 358
NAME/KEY: unsure
LOCATION: 359
NAME/KEY: unsure
LOCATION: 360
NAME/KEY: unsure
LOCATION: 361
NAME/KEY: unsure
LOCATION: 362
NAME/KEY: unsure
LOCATION: 363
NAME/KEY: unsure
LOCATION: 364
NAME/KEY: unsure
LOCATION: 365
NAME/KEY: unsure
LOCATION: 366
NAME/KEY: unsure
LOCATION: 367
NAME/KEY: unsure
LOCATION: 368
US-09-864-761-21746
```

Query Match 5.6%; Score 135.6; DB 10; Length 368;
Best Local Similarity 64.0%; Pred. No. 3.6e-26;
Matches 235; Conservative 0; Mismatches 105; Indels 27; Gaps 1;

```
472 TACAAGGAGACTATGTTTGGCCCATATTTACAGCGGCTGACATGCTAAATAATTCACAGA 531
12 TACAAGGAGAGCTACGGGCTCTCATATTTACAGCGCATGATGCTGCAGATCCCTAAA 71
532 CAACAAGACATCTCCATTTAAAGTGGCTCAATTCATGATCGACATCAAAAACATT 591
72 CAGACGCAAAAGCAAAATACAGTGCCTCAATTCGATCAACATTAATAATATTT 131
592 GATTCGGCAAAAGGATGATAGTACAGCAACTTAATAGATTAGATGTATGGATAGC 651
132 GAGTCTGCAAAAGGACTT-----GATGTGTGGACACG 164
652 TGGCCATCGCAAAAGCTGATGCTGCGCAAAATTTATCATGATATCATCGTCTCC 711
165 TGGCCGCTGCAAAAGCTGACGCAACACTGATCAACGCGCTATCAGCTGTGTTT 224
712 GCTTAGCAGTACCCCAAAACAGATGATATCTCATTCTATTCATCAAAA 771
225 GCTCTTGGGAGAGCCGAAAGAGCTGATGACACATCAATCTATTTATCAAAAG 284
772 GTCGGTATACATGATGACAGCTGGAATAATGCTGGAAGATTTTGAAGATATGAT 831
285 GTCGGCGACACATCAATGACAGCTGGAATAACGGGCGCGCTCTTTTAAAGACAGCAT 344
832 AAATTTG 838
```

```
Db 345 AAGTTCG 351
RESULT 10
US-09-974-300-735
Sequence 735, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M
APPLICANT: Clausen, Id Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 10085.500-US
CURRENT APPLICATION NUMBER: US/09/974.300
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/680.598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279.526
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 735
LENGTH: 882
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-735
```

Query Match 4.6%; Score 110.4; DB 10; Length 882;
Best Local Similarity 71.8%; Pred. No. 3.8e-19;
Matches 158; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

```
2190 GGATATCCGCCCTCAATFACCACCTTCTGTCGCCAAGTGAATAAGATCCGACAGG 2249
101 GATATCGTGACGCTTTCATTTGCAACCCCTGCAATGGAATAAGACCTCAAAA 160
2250 CCCATCTTTTGGGGA-AGTATCATTAATTAATCTTACAAAGACTATCCAGAT 2308
161 CCCGCTATTTCATGAGAAATATCATTAATTAATTAATTAATTAATTAATTAAT 220
2309 GGAATGATGAGATGCGGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2368
221 GGCATGGAACAGATGAGTGGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 280
2369 GGGATTGCCATCCGAGATATACCAATTAATAATGATGATC 2408
281 GGGGTGCCATTCGAAATATACGATATAAATGGCATC 320
```

RESULT 11
US-09-778-844-136
Sequence 136, Application US/09778844
Patent No. US20020150971A1
GENERAL INFORMATION:
APPLICANT: JOHANSEN, JEANETTE ELISABETH
APPLICANT: SCHALLING, MARTIN
TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
FILE REFERENCE: 030307/0195
CURRENT APPLICATION NUMBER: US/09/778.844
CURRENT FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 136
LENGTH: 1146
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: CPLA2-beta, EMBL
US-09-778-844-136

Query Match 3.9%; Score 93; DB 10; Length 1146;
Best Local Similarity 61.7%; Pred. No. 2.1e-14;

RESULT 12
US-09-864-761-5018
; Sequence 5018, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

1 APPLICANT: eenn, Sharon G.
 2 APPLICANT: Rank, David R.
 3 APPLICANT: Hanzel, David K.
 4 APPLICANT: Chen, Wensheng
 5 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 6 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 7 FILE REFERENCE: Aecmica-X-1
 8 CURRENT FILING DATE: 2001-05-23
 9 CURRENT FILING DATE: 2001-05-23
 10 PRIOR APPLICATION NUMBER: US 60/180,312
 11 PRIOR FILING DATE: 2000-02-04
 12 PRIOR APPLICATION NUMBER: US 60/207,456
 13 PRIOR FILING DATE: 2000-05-26
 14 PRIOR APPLICATION NUMBER: US 09/632,366
 15 PRIOR FILING DATE: 2000-08-03
 16 PRIOR APPLICATION NUMBER: GB 24263.6
 17 PRIOR FILING DATE: 2000-10-04
 18 PRIOR APPLICATION NUMBER: US 60/236,359
 19 PRIOR FILING DATE: 2000-09-27
 20 PRIOR APPLICATION NUMBER: PCT/US01/00666
 21 PRIOR FILING DATE: 2001-01-30
 22 PRIOR APPLICATION NUMBER: PCT/US01/00667
 23 PRIOR FILING DATE: 2001-01-30
 24 PRIOR APPLICATION NUMBER: PCT/US01/00668
 25 PRIOR FILING DATE: 2001-01-30
 26 PRIOR APPLICATION NUMBER: PCT/US01/00669
 27 PRIOR FILING DATE: 2001-01-30
 28 PRIOR APPLICATION NUMBER: PCT/US01/00665
 29 PRIOR FILING DATE: 2001-01-30
 30 PRIOR APPLICATION NUMBER: PCT/US01/00668
 31 PRIOR FILING DATE: 2001-01-30
 32 PRIOR APPLICATION NUMBER: PCT/US01/00663
 33 PRIOR FILING DATE: 2001-01-30
 34 PRIOR APPLICATION NUMBER: PCT/US01/00662
 35 PRIOR FILING DATE: 2001-01-30
 36 PRIOR APPLICATION NUMBER: PCT/US01/00661
 37 PRIOR FILING DATE: 2001-01-30
 38 PRIOR APPLICATION NUMBER: PCT/US01/00670
 39 PRIOR FILING DATE: 2001-01-30
 40 PRIOR APPLICATION NUMBER: US 60/234,687
 41 PRIOR FILING DATE: 2000-09-21
 42 PRIOR APPLICATION NUMBER: US 09/608,408
 43 PRIOR FILING DATE: 2000-06-30
 44 PRIOR APPLICATION NUMBER: US 09/774,203
 45 PRIOR FILING DATE: 2001-01-29

Query Match	3.38;	Score 80.2;	DB 10;	Length 376;
Best Local Similarity	58.5%;	Pred. No. 2.8e-11;		
Matches 176; Conservative	0;	Mismatches 98;	Indels 27;	Gaps 1

Db	100	ACACACTGCACCTCTGTGGCAGAGGAGACGACGACTCAAGCGCTTCGGGAAAGAAATAATACCAAAAA	159
QY	409	ACGCGTATATTTAGTAGACCGGACCGACCGCATTTTTCGCGAGCAAAATGAACACGGCG	468
QY	409	ACGCGTATATTTAGTAGACCGGACCGACCGCATTTTTCGCGAGCAAAATGAACACGGCG	468
QY	469	GACTACACAGAGACATATGCTTTTGCCCATATATACAGCGCGTGCATGCTAAAAATTCCA	528
Db	160	GCATACAAAGAAAACGTCACGGCGTCTCATATTAACGCCATGATATGTCGCAACATCCCT	219
QY	529	GCATACAAAGAACAGTCTTCATTTAAAGTGCCTCATTCATGCATGATGAGCAATCAAAAAC	588
Db	220	AAACGACGACGAAAACGAAAATAATACCAAGTGCCTCATTTGCATTCATCAACGATTTAAAAAT	279
QY	589	ATTGATTTCCGCAAAAAGGGTATGATAGTCAGGCAACTTAATAGATTAGATGTTGGGAT	648
Db	280	ATTGAACTCTGCAAAAAGG-----ACCTGATGTGTGGGAC	312
QY	649	AGCTGGCACTGCCAAAACGCTGATGTTACTGCGCAAAATTAATCAGATATACATGTC	708
Db	313	AGCTGGCCGCTGCAAAAACGCTGACGGAACAGTAGTGATATCAACAGCGGTATCAGTTCTG	372
QY	709	T-709	
Db	373	T-373	

RESULT 13
 US-09-995-587A-10
 Sequence 10, Application US/09995587A
 Patent No. US200217621A1
 GENERAL INFORMATION:
 APPLICANT: VAN HILUM, SACHA ADRIANDS FORKE TACO
 APPLICANT: VAN GEEL-SCHOUTEN, GERRIDINA HENDRIKA
 APPLICANT: DIJHUIZEN, IJLBERT
 APPLICANT: RAHNOU, HAKIM
 TITLE OF INVENTION: NOVEL FRUCTOSYLTRANSFERASES
 FILE REFERENCE: B043657-CIP
 CURRENT APPLICATION NUMBER: US/09/995,587A
 CURRENT FILING DATE: 2002-02-28
 PRIOR APPLICATION NUMBER: 09/604,958
 PRIOR FILING DATE: 2000-06-28
 PRIOR APPLICATION NUMBER: EPO 00201872.9
 PRIOR FILING DATE: 2000-05-25
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 10
 LENGTH: 4634
 TYPE: DNA
 ORGANISM: Lactobacillus reuteri
 FEATURE:

```

: NAME/KEY: CDS
: LOCATION: (1220)..(3598)
: FEATURE:
: NAME/KEY: RBS
: LOCATION: (1205)..(1210)
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (2702)..(2707)
: OTHER INFORMATION: a, c, t, g, other or unknown
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (3686)..(3698)
: OTHER INFORMATION: a, c, t, g, other or unknown
US-09-995-587A-10

Query Match
Best Local Similarity 2.5%; Score 61.2; DB 10; Length 4634;
Matches 181; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

OY 1047 TGATCATTAATCTGCTTTGATGCGGAGACGCTACAGTTTATTCATAATATTCAGCAATT 1106
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2299 TGTTCATATATATCAAGTTTGTGTAAGTGATGATTTCTACCAACTTATGAGACA 2358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1107 TATCGATGAAGGCAAGTGATTTTCAGTGATACCATCTTTAAGAGACCTCTACTATGT 1166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2359 ATTGCAACAGCGCAAGATCGTGAATAATGATGATTACTGCTTACGTCGACCCACAGTTGT 2418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1167 TGA---AGATGAAGGCGCATTAATATCTGCTTTGAGAGATTAAGTGAACAAGAGATG 1223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2419 TCAATTAAGAAAATGCTGATCTGTTATTCAGTAACTGAGTAACTGGACGAAGATTA 2478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1224 TTATCAAGGCGATCAGTCTTTCATATTAATTAACCTTACTATGAGGAGTACGCTTCTT 1283
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2479 CCAAGTGACGACCAATTTATTAATTTGGGCTAATGCTGCGCATMGATGCTTCAATAT 2538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1284 CCAAGATGAAAAAATAAAGCTCTCAAGCTCTTAAAAACAATGCTCTTTTACGGA 1343
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2539 TAAGAGTTCCTCAAGCTTTTGAATAATAGAAGATCCTGAATGCTGCTTGAAGCTAA 2598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1344 TGTGTCATTAGGACATTTGTAATTTGGCGGATGATATACAGTAAGTAAGTATTAAGACC 1403
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2599 TGTGTCACCTTGTGTTCTTAAGAGCTACTTAACATCAATGAAGTGAAGGTTGAAGAGT 2658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1404 AT 1405
    |||||
DB 2659 AT 2660
    |||||

RESULT 14
US-10-198-846-5829/c
: Sequence 5829, Application US/10198846
: Publication No. US2003009974A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yougyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steimann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: PRIOR FILING DATE: 2001-07-18
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: fastseq for Windows Version 4.0
: SEQ ID NO 5829
: LENGTH: 839
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature

```

```

: LOCATION: 3, 15, 188, 189, 190, 191, 192, 209, 222, 226, 227, 229,
: LOCATION: 241, 242, 244, 250, 257, 258, 259, 260, 261, 262, 263, 265,
: LOCATION: 281, 285, 291, 292, 293, 294, 295, 296, 298, 303, 305, 307,
: LOCATION: 308, 311, 312, 319, 323, 325, 329, 332, 339, 355, 357
: OTHER INFORMATION: n = A,T,C or G
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 358, 368, 369, 370, 374, 380, 386, 387, 388, 392, 401, 402,
: LOCATION: 403, 406, 409, 413, 414, 415, 426, 427, 439, 453, 460, 465,
: LOCATION: 466, 468, 470, 480, 482, 502, 503, 518, 524, 527, 534, 539,
: LOCATION: 547, 552, 557, 560, 569, 611, 612, 619, 627, 631, 640
: OTHER INFORMATION: n = A,T,C or G
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 644, 651, 658, 659, 672, 675, 676, 677, 679, 681, 691, 693,
: LOCATION: 698, 702, 720, 723, 724, 727, 732, 738, 742, 747, 748, 749,
: LOCATION: 750, 757, 758, 764, 769, 782, 783, 793, 800, 802, 817, 818,
: LOCATION: 821, 826, 828
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5829

Query Match
Best Local Similarity 2.2%; Score 52; DB 9; Length 839;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 1882 TTTTAAATAAAGGTTTGTGACACCTTTAGAGAAAAGAAATCTTGCGCCG 1941
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 610 TTTTTCCTCCAGAGGGGGGTTTTCCTTTTAAAGGAAAATGNGNGGCG 551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1942 GGCATTAACCATTCGCCAGAGAGCTACCAATAAAGAAATTTCCCTTTTAGCG 2001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 550 GGGNCCCCCGGNGGGGNGGNNTTNAAANNTTATATATATTTNNCCCCCGGGG 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 2002 GGGGACTCCTTTTCTTCTTCTTGGGTGGGATCTTGCCCCCACC 2049
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 490 GGGGGCCCTNTTTTAAANGNGGGGNTTTTNCCTCCCCC 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-09-995-587A-2
: Sequence 2, Application US/09995587A
: Patent No. US20020127681A1
: GENERAL INFORMATION:
: APPLICANT: VAN HIJUM, SACHA ADRIANUS FORKE TACO
: APPLICANT: VAN GEEL-SCHUTTEN, GERRITDIANA HENDRIKA
: APPLICANT: DIJKHUIZEN, LOBBERT
: APPLICANT: RAHOUTI, HAKIM
: TITLE OF INVENTION: NOVEL FRUCTOSYLTRANSFERASES
: FILE REFERENCE: BO43667-CIP
: CURRENT APPLICATION NUMBER: US/09/995,587A
: CURRENT FILING DATE: 2002-02-28
: PRIOR APPLICATION NUMBER: 09/604,958
: PRIOR FILING DATE: 2000-06-28
: PRIOR APPLICATION NUMBER: EPO 00201872.9
: PRIOR FILING DATE: 2000-05-25
: NUMBER OF SEQ ID NOS: 40
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 2367
: TYPE: DNA
: ORGANISM: Lactobacillus reuteri
US-09-995-587A-2

Query Match
Best Local Similarity 2.2%; Score 52; DB 10; Length 2367;
Matches 589; Conservative 0; Mismatches 535; Indels 96; Gaps 12;

OY 536 AAACAGCTCCATTAATTAAAGTCCCATTCATGATCAGCATCAAGCAATGATT 595
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 674 AACAGATGCTGCGTACACAGTTCATTTTAAAGCAAGTAAATCAAAATATGCTG 733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 596 CGGCAAAAGGATATAGATGAGGACACTTAATAGA---TTTAGATGATGAGTACT 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

D	b	734	CCGGTACAACTAAAGATSCAACAACATACATACCTATTGGAACCTTTAGATGATAGGATTCAT	793
O	y	653	GGCCACTGCAAAACGCT--GATGGTACTCGCGCAAAATTATCATGATATCACATCGTCT	709
D	b	794	GGCCAGTTCAAGATGTTTCGGACAGACAAAGTGGTAATTGGAAAGGCGATCAACTTGCA	853
O	y	710	CCGGTTTACGAGTGAGCAAAAACAGTGAATGATGCCACTCATTTATTCATCAAA	769
D	b	854	TGCGAATGATGGGAATTCCAACCAAA--ATGATAATCATATCTGCTTTATATATA	910
O	y	770	AAGTCGGTATACATCATGATGACAGCTGCGAAAAATGCTGGAAGAGTATTCAGAAATATGG	829
D	b	911	AGTATGGTGAATATGAATTAAGTCAATTGGAAAGATGTGCTGTCCAATTTTGGCTATATTT	970
O	y	830	ATAAATTTGTTCCAAATGATCCGTAATCTTAAATATCAACACAGAGTGGTCAGGTTCTG	889
D	b	971	CTACCGCGGTT-----TCACAGAAATGCTCAGGATCAG	1003
O	y	890	CTACTTTACCAAGATGGCCAAATGCCCTTATTTCTATACAGATTACTCAGTAAATCTG	949
D	b	1004	CTGTTTGAACAGTATACCTATCTATCCAAATTAATTTATACA-----AGGGTAG	1051
O	y	950	AAGATGGTGAACCGGTGCTGGTAAACCAATCATTTTCAACTGCTCAAGTAACTAATCC	1009
D	b	1052	ACAGCTCGATACAAATACCAATACCAATCATCAAAAATTTGCTAGCGCTACTTTATTTAACTG	1111
O	y	1010	AGCCGATGACGTACACTTAAAGTCGATGATCTGATCATTAATCTGCTTTGATG	1069
D	b	1112	ATAAATATGAAATATGATACCTCCCTCAGTACGAAATGACTATA-----TTGATATTG	1165
O	y	1070	GGCGAGACGGTACAGTTTATCAAAATATTCAGCAATTTATCGATGAGGCAAGTGATTT	1129
D	b	1166	AAGGTGATGGCTATTACTACCAAACTTATGATCAATGGAAAGCTACTAACCAAAAG--TG	1222
O	y	1130	CAGGTGATTAACCTACTTTAAGAGACCCCTCATGTTGTAAGATTAAGGCCCATTAATATC	1189
D	b	1223	CCGATTAATATGCAATGCGTGATGCTCATGTAATTTGAAGATGATATGATGCGTACC	1282
O	y	1190	TTGCTTTTGAAGCAATACCTGGAACAACAGATGTTATCAAGGGCATCAGTCTTTCATA	1249
D	b	1283	TTGTTTTTGAAGCAAGTACTGG--TTTGAAATTTATCAAGGCGAGGACCAAAATTTATA	1339
O	y	1250	ATTAAGCTTACTATGGGGGAGTACAGCTCTTCTCCAGAAATGAAAAAATTAATCTGCTC	1309
D	b	1340	ACTGTTAAATTAATGCGGAGATGATCGCATTTAATATCAAGGCTTATTTGAATTCCTT	1399
O	y	1310	AAAGTCCATAAAAACAAAAT--GCTTCTTACGGAATGCGCAATTAAGCAATTTGATGAT	1366
D	b	1400	CCATGATGATTAATTAAGAGTGGGCACTTGGCTAATGACGCTATCGTATCTCTCAAC	1459
O	y	1367	TGCGCGATGACTA-----TACAGTGAAGAAGTATTGAACAACATTAAGTCGATCAA	1417
D	b	1460	TAAATTAAGAGCAAGAAAGATCTTAAGGCGCAGAGTATATCTCACCATTAATTTCTGCAC	1519
O	y	1418	ACACAGTATGAGATGANGATGCAAGCGGCATATATTTAAATGAATTAATTAATGATGATC	1477
D	b	1520	CAATGTGTAAGCAATGAATTTAGGACCAAAATGATTAATTTAGGTAAATTAATTAATCT	1579
O	y	1478	TATTCACGAGATTCAGAGGATCCAAAATGACAGAGATGGAATTAACAGCAAAATGTT	1537
D	b	1580	TATTTGCCCGCTACCCGTTTAAATGAGAGAGTAATGATGATGCTTGGATGATCTAATTT	1639
O	y	1538	AT-----ATGCTAGGGCCCGAGGCGACCTCTTAATGAGCCCAAC	1576
D	b	1640	ATGCCGTTGGATATATGTTGCAATGTGGCGGATATGTTGGCGATATGCTAATCTGATCTTT	1699
O	y	1577	ACAACCCGATTAATGAACCTGGAATCTGATTTGAACATGAATCTTGAACCTGCTGATCTCA	1636
D	b	1700	ATAAGCCATTAATAAGATTTCTGAGATGCTTGAATCTGCTGCTGCTGCAAACTGGCGGA	1759
O	y	1637	CACACACTTACTCTATTTGGGCTATCCCGCACCCCTG--AAGTAATATATGTGCTACTCA	1693
D	b	1760	CAGCACTTATTCATATTAATGCTCTCCCGTGGCGGAAAGATGACCACCAATATTTAGCTTA	1819

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 02:52:36 ; Search time 111.688 seconds
(without alignments)
6611.951 Million cell updates/sec

Title: US-09-986-682b-5

Perfect score: 2408
Sequence: 1 CCGCGAATAACTAGATTC.....TACCATTAATAATGTGATC 2408

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2408	100.0	2408	2	US-08-870-827-5
2	2408	100.0	2408	4	US-09-317-179-5
3	1365	56.7	1365	2	US-08-870-827-4
4	1365	56.7	1365	4	US-09-317-179-4
5	607.2	25.2	10317	3	US-09-058-746-1
6	607.2	25.2	10317	4	US-09-438-142-1
7	553.8	23.0	3305	3	US-09-068-043-1
8	44.8	1.9	7218	1	US-08-232-463-14
9	39.8	1.7	1662	1	US-08-671-947-1
10	38	1.6	807	4	US-09-276-531-82
11	37.6	1.5	1434	4	US-09-134-001C-650
12	36.6	1.5	7101	1	US-08-480-604A-9
13	36.6	1.5	7101	2	US-08-405-496A-9
14	36.6	1.5	7101	4	US-08-915-136-9
15	36.6	1.5	7101	4	US-08-957-310-9
16	36.4	1.5	3095	6	5231168-1
17	35.8	1.5	381	4	US-09-134-001C-1861
18	35.4	1.5	849	4	US-09-648-520E-48
19	35.4	1.5	3317	4	US-09-193-562D-1
20	35.4	1.5	6854	4	US-08-961-527-200
21	34.8	1.4	62	3	US-09-068-043-5
22	34.2	1.4	3360	1	US-07-712-833A-1
23	34	1.4	1467	5	PCT-US93-05944-1
24	34	1.4	2817	5	US-09-379-523-4
25	34	1.4	3356	4	US-08-446-137B-1
26	33.8	1.4	3279	4	US-08-947-823-1
27	33.8	1.4	51952	3	Sequence 1, Appl1

C 28	33.6	1.4	1365	2	US-08-870-827-4	Sequence 4, Appl1
C 29	33.6	1.4	1365	4	US-09-317-179-4	Sequence 4, Appl1
C 30	33.6	1.4	2408	2	US-08-870-827-5	Sequence 5, Appl1
C 31	33.6	1.4	2408	4	US-09-317-179-5	Sequence 5, Appl1
C 32	33.6	1.4	3713	1	US-08-100-709-1	Sequence 1, Appl1
C 33	33.6	1.4	3713	1	US-08-176-865-1	Sequence 1, Appl1
C 34	33.6	1.4	3713	1	US-08-474-038-1	Sequence 1, Appl1
C 35	33.6	1.4	3713	2	US-08-779-046-1	Sequence 1, Appl1
C 36	33.6	1.4	3713	2	US-08-881-340-1	Sequence 1, Appl1
C 37	33.4	1.4	1566	4	US-09-134-001C-2802	Sequence 2802, Ap
C 38	33.4	1.4	2022	4	US-08-879-098-1	Sequence 1643, Ap
C 39	33.2	1.4	900	4	US-08-631-548-1	Sequence 1, Appl1
C 40	33.2	1.4	900	4	US-09-631-548-1	Sequence 1, Appl1
C 41	33.2	1.4	3891	1	US-08-480-604A-27	Sequence 27, Appl1
C 42	33.2	1.4	3891	2	US-08-405-496A-27	Sequence 27, Appl1
C 43	33.2	1.4	3891	4	US-08-915-136-27	Sequence 27, Appl1
C 44	33	1.4	699	4	US-08-936-165A-233	Sequence 23, App
C 45	33	1.4	2047	4	US-08-836-261A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-870-827-5
Sequence 5, Application US/08870827
Patent No. 5962297
GENERAL INFORMATION:
APPLICANT: Tsusaki et al.
TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE
NUMBER OF SEQUENCES: 6
CORRESPONDING ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,827
FILING DATE: 06-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 170,630/1996
FILING DATE: 10-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2408 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus sp.
INDIVIDUAL ISOLATE: V230 (FERM BP-5054)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..360
IDENTIFICATION METHOD: E
NAME/KEY: signal peptide

LOCATION: 361..456
IDENTIFICATION METHOD: S
NAME/KEY: mat peptide
LOCATION: 457..1821
IDENTIFICATION METHOD: S
NAME/KEY: 3'UTR
LOCATION: 1822..2408
IDENTIFICATION METHOD: E
US-08-870-827-5

Query Match 100.0%; Score 2408; DB 2; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGGAAAATACTAGATTCCTCAATGGCCAGACTTCCAGTGTGTAGAGAGAGTTGG 60
Db 1 CGGGGAAAATACTAGATTCCTCAATGGCCAGACTTCCAGTGTGTAGAGAGAGTTGG 60
QY 61 GACTGCCAATGCACTGTGCGTAAGAAAACAGCTTACTCATGAGCAATTACTAGAGAT 120
Db 61 GACTGCCAATGCACTGTGCGTAAGAAAACAGCTTACTCATGAGCAATTACTAGAGAT 120
QY 121 TTCAAAGTCTGGGATTAAGCCCAAGTCCACTTTGAATTAACCTTTGAGCCTGTGTG 180
Db 121 TTCAAAGTCTGGGATTAAGCCCAAGTCCACTTTGAATTAACCTTTGAGCCTGTGTG 180
QY 181 GGGGCTTTTGTATTATTATTATTTCAACTGCAAGTGTCCATCCCTATATCAATTAA 240
Db 181 GGGGCTTTTGTATTATTATTATTTCAACTGCAAGTGTCCATCCCTATATCAATTAA 240
QY 181 GGGGCTTTTGTATTATTATTATTTCAACTGCAAGTGTCCATCCCTATATCAATTAA 240
Db 181 GGGGCTTTTGTATTATTATTATTTCAACTGCAAGTGTCCATCCCTATATCAATTAA 240
QY 241 GAGCAAAATCTAATCAATCCATGCCATCCCAATTAACCTGCTCTCTATCTTTTAA 300
Db 241 GAGCAAAATCTAATCAATCCATGCCATCCCAATTAACCTGCTCTCTATCTTTTAA 300
QY 241 GAGCAAAATCTAATCAATCCATGCCATCCCAATTAACCTGCTCTCTATCTTTTAA 300
Db 241 GAGCAAAATCTAATCAATCCATGCCATCCCAATTAACCTGCTCTCTATCTTTTAA 300
QY 301 TTAATTAAGAACTATCAAGACTTTCTTATCAATTCATATATCCAGAGAGAGAGAG 360
Db 301 TTAATTAAGAACTATCAAGACTTTCTTATCAATTCATATATCCAGAGAGAGAGAG 360
QY 301 TTAATTAAGAACTATCAAGACTTTCTTATCAATTCATATATCCAGAGAGAGAGAG 360
Db 301 TTAATTAAGAACTATCAAGACTTTCTTATCAATTCATATATCCAGAGAGAGAGAG 360
QY 361 ATGAACCTTCAAAAGATTGGCGAAAAAAGCACTGCCGTACCTTCAGAGACTGTATTTA 420
Db 361 ATGAACCTTCAAAAGATTGGCGAAAAAAGCACTGCCGTACCTTCAGAGACTGTATTTA 420
QY 421 GTAGAGCGGAGGACCGCATATTTTGGGAGCAAAATGAACAGCGGGAGCTACAGAGAA 480
Db 421 GTAGAGCGGAGGACCGCATATTTTGGGAGCAAAATGAACAGCGGGAGCTACAGAGAA 480
QY 481 GACTATGTTTGGCCATATTAACAGCGGCTGACATGCTAAATTTCCAGAGCAAAAC 540
Db 481 GACTATGTTTGGCCATATTAACAGCGGCTGACATGCTAAATTTCCAGAGCAAAAC 540
QY 541 AGTCCCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAACATTGATTGCGCA 600
Db 541 AGTCCCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAACATTGATTGCGCA 600
QY 541 AGTCCCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAACATTGATTGCGCA 600
Db 541 AGTCCCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAACATTGATTGCGCA 600
QY 601 AAAGGATATGATAGTACAGCACTTAATAGATTTAGATGTATGATGATGATGATGAT 660
Db 601 AAAGGATATGATAGTACAGCACTTAATAGATTTAGATGTATGATGATGATGATGAT 660
QY 601 AAAGGATATGATAGTACAGCACTTAATAGATTTAGATGTATGATGATGATGATGAT 660
Db 601 AAAGGATATGATAGTACAGCACTTAATAGATTTAGATGTATGATGATGATGATGAT 660
QY 661 CAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 CAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GGTGACCCCAAAAACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 GGTGACCCCAAAAACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

QY 901 AAAGTGGCCCAAGTCCGTTTATCTATACAGATTAAGTATCCGAAGATGATGGA 960
Db 901 AAAGTGGCCCAAGTCCGTTTATCTATACAGATTAAGTATCCGAAGATGATGGA 960
QY 961 ACCGATGCTGTAAACCAATATTTCACTGCTCAAGTAACTTATCCAGCCGGATGCA 1020
Db 961 ACCGATGCTGTAAACCAATATTTCACTGCTCAAGTAACTTATCCAGCCGGATGCA 1020
QY 1021 GCTACCTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1021 GCTACCTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACAGTTTATCAAAATATTTCAAGCAATTTATGATGATGATGATGATGATGATGATGAT 1140
Db 1081 ACAGTTTATCAAAATATTTCAAGCAATTTATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 CATACCTTAAAGACCCCTCACTATGTTGAAGATTAAGGCCATAAATATCTTGTGAA 1200
Db 1141 CATACCTTAAAGACCCCTCACTATGTTGAAGATTAAGGCCATAAATATCTTGTGAA 1200
QY 1201 GCGAATACGTGAACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1201 GCGAATACGTGAACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 TATGCGGAGTGAAGTGTCTTCTCCAGATTAAGAAAAATTAACCTGTTCAAGCTTAA 1320
Db 1261 TATGCGGAGTGAAGTGTCTTCTCCAGATTAAGAAAAATTAACCTGTTCAAGCTTAA 1320
QY 1321 AAACAAATGCTTCTTAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 AAACAAATGCTTCTTAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 ACAGTGAAGTGTATGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db 1381 ACAGTGAAGTGTATGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 CGCGCAATATATTTAAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1500
Db 1441 CGCGCAATATATTTAAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1500
QY 1501 AAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 AAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 TCTTAAATGAGCCACACACACACACACACACACACACACACACACACACACACACACAC 1620
Db 1561 TCTTAAATGAGCCACACACACACACACACACACACACACACACACACACACACACACAC 1620
QY 1621 GACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 GACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 CTGCGGAGCAAGTGTGAGTAAATGAAGGCTGACACATCTGAGAGGAGAAATAGT 1800
Db 1741 CTGCGGAGCAAGTGTGAGTAAATGAAGGCTGACACATCTGAGAGGAGAAATAGT 1800
QY 1801 TCCGAGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 TCCGAGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 GATCTTACCTTTTCCCAACCTTTTAAATGAAGGCTTGTACCACTTTAGAA 1920
Db 1861 GATCTTACCTTTTCCCAACCTTTTAAATGAAGGCTTGTACCACTTTAGAA 1920
QY 1921 GAAAAAAGATTCCTTGGCCCGGCAATTAACCATTCGCAAGAGCTACCAATTAATA 1980
Db 1921 GAAAAAAGATTCCTTGGCCCGGCAATTAACCATTCGCAAGAGCTACCAATTAATA 1980

1981 AGAATTTTCCCTTTTATAGGGGGGAGTCTTTTCTTATCTTGGGGTGGGATTTG 2040
1981 AGAATTTTCCCTTTTATAGGGGGGAGTCTTTTCTTATCTTGGGGTGGGATTTG 2040
2041 GCGCCCAACGAGGAGCTTTTATTTATTTATGATGATCTGGGATTTTCCGCGCTA 2100
2041 GCGCCCAACGAGGAGCTTTTATTTATTTATGATGATCTGGGATTTTCCGCGCTA 2100
2101 CTTATTTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2160
2101 CTTATTTTCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2160
2161 ACAGTTTACCAACACCTGCTGGAGGAGGATGATGCGCCCTCAATACCACTTTCTGTC 2220
2161 ACAGTTTACCAACACCTGCTGGAGGAGGATGATGCGCCCTCAATACCACTTTCTGTC 2220
2221 CCGCAATGGAATAATGATCCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280
2221 CCGCAATGGAATAATGATCCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280
2281 TTACCTTTACCAACAAAGCTATCCAGATGGAATGTAACGAGATGCGGCGGCGGCGGCGG 2340
2281 TTACCTTTACCAACAAAGCTATCCAGATGGAATGTAACGAGATGCGGCGGCGGCGGCGG 2340
2341 CGAAGATTTATGCTGATGAGAGGAGGAGGATGCGGCGGCGGCGGCGGCGGCGGCGG 2400
2341 CGAAGATTTATGCTGATGAGAGGAGGAGGATGCGGCGGCGGCGGCGGCGGCGGCGG 2400
2401 TGGTGATC 2408
2401 TGGTGATC 2408

RESULT 2
US-09-317-179-5
Sequence 5, Application US/09317179
Patent No. 6383769
GENERAL INFORMATION:
APPLICANT: Tsusaki et al.
TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE ACTIVITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/317,179
FILING DATE: 24-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/870,827
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2408 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Bacillus sp.
INDIVIDUAL ISOLATE: V230 (FERM BP-5054)
NAME/KEY: signal peptide
LOCATION: 1..360
IDENTIFICATION METHOD: E
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-317-179-5
Query Match 100.0%; Score 2408; DB 4; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCGGGAATAATCTAGATTCCAAATGGCCAGACTTCCAGTTGGGTAAAGAGAGTTG 60
1 CCGGGAATAATCTAGATTCCAAATGGCCAGACTTCCAGTTGGGTAAAGAGAGTTG 60
61 GACTGCCAATGCGAGTGGGTGAAGAAACAGCTTACTGATGAGCAATTACTAGAGAT 120
61 GACTGCCAATGCGAGTGGGTGAAGAAACAGCTTACTGATGAGCAATTACTAGAGAT 120
61 GACTGCCAATGCGAGTGGGTGAAGAAACAGCTTACTGATGAGCAATTACTAGAGAT 120
121 TTCAAAAGTCTGGGATTAAGCCAAAGTCCACTTTGAAATTAACCTTTGAGCTGTGTG 180
121 TTCAAAAGTCTGGGATTAAGCCAAAGTCCACTTTGAAATTAACCTTTGAGCTGTGTG 180
121 TTCAAAAGTCTGGGATTAAGCCAAAGTCCACTTTGAAATTAACCTTTGAGCTGTGTG 180
181 GGGGCTTTTGTGTTTATTTATTTCAACTGCAAGTGTCCATCCCTATATCAATTTAA 240
181 GGGGCTTTTGTGTTTATTTATTTCAACTGCAAGTGTCCATCCCTATATCAATTTAA 240
241 GACGAATTTTAAATCAATCCATGCCATGCCAATTAATGCTGCTCTATATCTTTAA 300
241 GACGAATTTTAAATCAATCCATGCCATGCCAATTAATGCTGCTCTATATCTTTAA 300
241 GACGAATTTTAAATCAATCCATGCCATGCCAATTAATGCTGCTCTATATCTTTAA 300
301 TTAATTAAGAACTATCAAGAGCTTTCTATCAAAATTCATATATCAAGAGAGGAGAG 360
301 TTAATTAAGAACTATCAAGAGCTTTCTATCAAAATTCATATATCAAGAGAGGAGAG 360
361 ATGAATTTCAAAAGATTTGGCGAAAAGACAGCTGCGGTACCTTACGATGCTATATTA 420
361 ATGAATTTCAAAAGATTTGGCGAAAAGACAGCTGCGGTACCTTACGATGCTATATTA 420
421 GTAGAGCGGAGCGGAGCGCATATTTTGGCGAGAAATGAAGAGGGGAGCTACAGGAA 480
421 GTAGAGCGGAGCGGAGCGCATATTTTGGCGAGAAATGAAGAGGGGAGCTACAGGAA 480
481 GACTATGTTTTGCCATATTACACGCGCTGACATGCTAATAAATTCAGACAAACAAAC 540
481 GACTATGTTTTGCCATATTACACGCGCTGACATGCTAATAAATTCAGACAAACAAAC 540
541 AGTCTCAATTTAAAGTGCCTCAATTCATATGATGAGCAATCAAAACATTGATTGGCA 600
541 AGTCTCAATTTAAAGTGCCTCAATTCATATGATGAGCAATCAAAACATTGATTGGCA 600
601 AAAGGTTGATTAAGTCAAGCACTTAATAGATTAGATGAGGAGAGTGGCGACATG 660
601 AAAGGTTGATTAAGTCAAGCACTTAATAGATTAGATGAGGAGAGTGGCGACATG 660
661 CAAAGCGTGATGCTAGTGGCGCAAAATTTATGATGATATCAATGCTCCGCTTTAGCA 720
661 CAAAGCGTGATGCTAGTGGCGCAAAATTTATGATGATATCAATGCTCCGCTTTAGCA 720
721 GGTGACCCCAAAACAGATGATGATCACTCACTTATTTCTATCAAAAGTGGTGAT 780
721 GGTGACCCCAAAACAGATGATGATCACTCACTTATTTCTATCAAAAGTGGTGAT 780
781 ACATGATTTGACAGTGAAGAAATGCTGAGAGAGATTTGAAGATGAGATTAATTTGTT 840
781 ACATGATTTGACAGTGAAGAAATGCTGAGAGAGATTTGAAGATGAGATTAATTTGTT 840
841 CCAATGATCGGTATCTTAATATCAAAACAGAGAGTGGTGAGTCTGCTACTTTAAC 900

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: Linear
US-08-870-827-4

Query Match 56.7%; Score 1365; DB 2; Length 1365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

457 ATGAACGGGGGACTACAGAGACATATGTTGGCCATATATACACGGCGTACATG 516
|||||
1 ATGAACGGGGGACTACAGAGAGACATATGTTGGCCATATATACACGGCGTACATG 60
517 CTAATAATCCAGAGACAAACAGTCTCAATTTAAAGTCCCAATTAATGATCA 576
|||||
61 CTAATAATCCAGAGACAAACAGTCTCAATTTAAAGTCCCAATTAATGATCA 120
577 GCAATCAAAAACATTGATGGGCAAAAGGTAATGATAGCAGCCACTTAATGATTTA 636
|||||
121 GCAATCAAAAACATTGATGGGCAAAAGGTAATGATAGCAGCCACTTAATGATTTA 180
637 GATGATGGGATAGTGGCCACTGCAAAAGCTGATGTACTGGCGCAATTAATGATGA 696
|||||
181 GATGATGGGATAGTGGCCACTGCAAAAGCTGATGTACTGGCGCAATTAATGATGA 240
697 TATCACAATCTCTCCGCTTTAGCAGGTGACCAAAAACAGTATGATCTCCACTTCA 756
|||||
241 TATCACAATCTCTCCGCTTTAGCAGGTGACCAAAAACAGTATGATCTCCACTTCA 300
757 TTATCTATCAAAAAGTGGTGAATACATGATGACAGCTGGAAAAATGCTGAAGAGTA 816
|||||
301 TTATCTATCAAAAAGTGGTGAATACATGATGACAGCTGGAAAAATGCTGAAGAGTA 360
817 TTGAAGATATGATGAATTTGTTCCAAATGATCCGTTCTTAATATCAACACAGAG 876
|||||
361 TTGAAGATATGATGAATTTGTTCCAAATGATCCGTTCTTAATATCAACACAGAG 420
877 TGGTCAGGTTCTGCTACTTTAAACCAAGATGGCCAAAGTCCGTTATTTCTATACAGATTAC 936
|||||
421 TGGTCAGGTTCTGCTACTTTAAACCAAGATGGCCAAAGTCCGTTATTTCTATACAGATTAC 480
937 TCAGGTAATCTCTGAAGATGTTGAACCGGCTCTGCTGAACCAAAATCAATTCATGCTCAA 996
|||||
481 TCAGGTAATCTCTGAAGATGTTGAACCGGCTCTGCTGAACCAAAATCAATTCATGCTCAA 540
997 GTAACCTATCCGAGCGGATGACGCTACCTTAAGTCGATGAGATTCGATCATATAA 1056
|||||
541 GTAACCTATCCGAGCGGATGACGCTACCTTAAGTCGATGAGATTCGATCATATAA 600
1057 TCTGCTCTTTGATGGCGAGAGCGTACAGTTTATCAAAATATTCAGCAATTTATCGATGAA 1116
|||||
601 TCTGCTCTTTGATGGCGAGAGCGTACAGTTTATCAAAATATTCAGCAATTTATCGATGAA 660
1117 GGCAGATGATTTGAGTGATGATACCATATCTTTAAGAGACCTCACTATGTTGAAGATTAAG 1176
|||||
661 GGCAGATGATTTGAGTGATGATACCATATCTTTAAGAGACCTCACTATGTTGAAGATTAAG 720
1177 GGCATATAATCTGCTCTTTGAGCGCAATACTGGAACAACAGATGTTATCAAGGCAT 1236
|||||
721 GGCATATAATCTGCTCTTTGAGCGCAATACTGGAACAACAGATGTTATCAAGGCAT 780
1237 CAGTCTTTCAATATAAGCTTACTATGCGGAAGTGCCTCTTCTCCAGAAATGAAAAA 1296
|||||
781 CAGTCTTTCAATATAAGCTTACTATGCGGAAGTGCCTCTTCTCCAGAAATGAAAAA 840
1297 AATAAATGCTTCAAAAGTCTTAAAAAACAAATGCTTCTTTAGCGAATGTCATTAAGC 1356
|||||
841 AATAAATGCTTCAAAAGTCTTAAAAAACAAATGCTTCTTTAGCGAATGTCATTAAGC 900
1357 ATGTTGAATGGCGGATGATATACAGTGAAGTGTATGAACCATTTAGTGCATCA 1416
|||||
901 ATGTTGAATGGCGGATGATATACAGTGAAGTGTATGAACCATTTAGTGCATCA 960

1417 AACACAGTAGAGATGAAGTCGACGCCGCAATATATTTAAATGAATATAATGCTAT 1476
|||||
961 AACACAGTAGAGATGAAGTCGACGCCGCAATATATTTAAATGAATATAATGCTAT 1020
1477 CTATTCACGAGATTCAAGAGATTCCAAATAGCAGAGTATGGAATTAACGACAAAGATTT 1536
|||||
1021 CTATTCACGAGATTCAAGAGATTCCAAATAGCAGAGTATGGAATTAACGACAAAGATTT 1080
1537 TATATGCTAAGGCGCGGAGGCGACCTCTTAATATGCGCCACACACACCATTAATGAACCT 1596
|||||
1081 TATATGCTAAGGCGCGGAGGCGACCTCTTAATATGCGCCACACACCATTAATGAACCT 1140
1597 GCACTTGTATGAACATCAATCTTGACCCCTGCTGATCTCACACACACTTACTCTATTGC 1656
|||||
1141 GCACTTGTATGAACATCAATCTTGACCCCTGCTGATCTCACACACACTTACTCTATTGC 1200
1657 GGTATCCGACCCCTGGAAGTATATATGCTGCTACACAGTTATATGACAAATAGAGGC 1716
|||||
1201 GGTATCCGACCCCTGGAAGTATATATGCTGCTACACAGTTATATGACAAATAGAGGC 1260
1717 TTCTATCCGAACATCACTCTCACCTCGGCGGACAGCTTGGGTTAATTAAGGCTCT 1776
|||||
1281 TTCTATCCGAACATCACTCTCACCTCGGCGGACAGCTTGGGTTAATTAAGGCTCT 1320
1777 GACACATCTGAGAGCAAAATAGTCCGAGACAGACAAATTCOCA 1821
|||||
1321 GACACATCTGAGAGCAAAATAGTCCGAGAGACAAATTCOCA 1365

RESULT 4
US-09-317-179-4
; Sequence 4, Application us/09317179
; Patent No. 6383769
GENERAL INFORMATION:
; APPLICANT: Tsusaki et al.
; TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE ACTIVITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
; ADDRESS: BROWDY AND NEWMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/317,179
; FILING DATE: 24-May-1999
; CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/870,827
; FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TSUSAKI=2
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-317-179-4

Query Match 56.7%; Score 1365; DB 4; Length 1365;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 457 ATGAACACGGGGAGCTACAGAGAGACTATGTTTGGCCATATTACACGGCGTACATG 516
DB 1 ATGAACACGGGGAGCTACAGAGAGACTATGTTTGGCCATATTACACGGCGTACATG 60
OY 517 CTAATAATTCAGAGCAACAAACAGTCTCAATTAAGTGCCCAATTCATGCAATCA 576
DB 61 CTAATAATTCAGAGCAACAAACAGTCTCAATTAAGTGCCCAATTCATGCAATCA 120
OY 577 GCAATCAAAACATTCGATTCGGCAAAAGGATGATGATGATGATGATGATGATGAT 636
DB 121 GCAATCAAAACATTCGATTCGGCAAAAGGATGATGATGATGATGATGATGATGAT 180
OY 637 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
DB 181 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
OY 697 TATCATCATGCTCCGCTTTAGCAGGTGACCCCAAAACAGTATGATGATGATGATGAT 756
DB 241 TATCATCATGCTCCGCTTTAGCAGGTGACCCCAAAACAGTATGATGATGATGATGAT 300
OY 757 TTATTCATCAAAAAGTGGTGATACATGATGATGATGATGATGATGATGATGATGAT 816
DB 301 TTATTCATCAAAAAGTGGTGATACATGATGATGATGATGATGATGATGATGATGAT 360
OY 817 TTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
DB 361 TTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
OY 877 TGGTGAAGTGTGCTACTTTTAAACAAAGATGCGCAAGTCCGTTTATTCATGACATTA 936
DB 421 TGGTGAAGTGTGCTACTTTTAAACAAAGATGCGCAAGTCCGTTTATTCATGACATTA 480
OY 937 TCAGTAAATCCTGAAAGATGATGGAACCGGTGCTGTAACCAATCATTTCAACTGCTCA 996
DB 481 TCAGTAAATCCTGAAAGATGATGGAACCGGTGCTGTAACCAATCATTTCAACTGCTCA 540
OY 997 GTAACCTATCCAGCGGATGACCTACATTAAGTGAAGGATGATGATGATGATGATGAT 1056
DB 541 GTAACCTATCCAGCGGATGACCTACATTAAGTGAAGGATGATGATGATGATGATGAT 600
OY 1057 TCTGCTTTGATGCGGAGACGCTACAGTTTATCAAAATATTACCAATTTTTCATGATA 1116
DB 601 TCTGCTTTGATGCGGAGACGCTACAGTTTATCAAAATATTACCAATTTTTCATGATA 660
OY 1117 GGCAGATGATTCAGTGATTAACCATTTTAAGAGACCTGACTATGTTGAAGATGATG 1176
DB 661 GGCAGATGATTCAGTGATTAACCATTTTAAGAGACCTGACTATGTTGAAGATGATG 720
OY 1177 GGCATTAATATCTGTCTTTGAAGCGAATCTGGAACAGATGATGATGATGATGATGAT 1236
DB 721 GGCATTAATATCTGTCTTTGAAGCGAATCTGGAACAGATGATGATGATGATGATGAT 780
OY 1237 CAGTCTTCAATTAATAGCTTACTATGCGGAGTGAAGTCTTCTTTCAGAGATGAAAA 1296
DB 781 CAGTCTTCAATTAATAGCTTACTATGCGGAGTGAAGTCTTCTTTCAGAGATGAAAA 840
OY 1297 AATAAAGCTTCAAAAGTCTTAAACAAATTTGCTTTTGAAGCAATGTTGATGATGATG 1356
DB 841 AATAAAGCTTCAAAAGTCTTAAACAAATTTGCTTTTGAAGCAATGTTGATGATGATG 900
OY 1357 ATTGTTGATGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1416
DB 901 ATTGTTGATGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
OY 1417 AACACAGTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1476
DB 961 AACACAGTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
OY 1477 CTATTCACGATTCAGAGAGATCAAAATGACGAGTATGATGATGATGATGATGATGAT 1536
  
```

```

DB 1021 CTATTCACGATTCAGAGAGATCCAAATGACGAGTATGATGATGATGATGATGATGAT 1080
OY 1537 TATATGATGAGGCGCGGAGGAGCTCCTTAATATGATGATGATGATGATGATGATGATGAT 1596
DB 1081 TATATGATGAGGCGCGGAGGAGCTCCTTAATATGATGATGATGATGATGATGATGATGAT 1140
OY 1597 GGACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
DB 1141 GGACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
OY 1657 GGTATCCGCAACCTGGAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1716
DB 1201 GGTATCCGCAACCTGGAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
OY 1717 TTTATTCAGAGACATCATCTCTACCTGCGGAGCAAGTGTGGGTTAATTAAGGGTCT 1776
DB 1261 TTTATTCAGAGACATCATCTCTACCTGCGGAGCAAGTGTGGGTTAATTAAGGGTCT 1320
OY 1777 GACACATCTGGAGAGAAATAGTTCGAGACAGACATTCACA 1821
DB 1321 GACACATCTGGAGAGAAATAGTTCGAGACAGACATTCCCA 1365
  
```

```

RESULT 5
US-09-058-746-1/c
: Sequence 1, Application US/09058746
: Patent No. 6022716
: GENERAL INFORMATION:
: APPLICANT: Ilya Chumakov
: APPLICANT: Hiroaki Tanaka
: TITLE OF INVENTION: High throughput DNA sequencing vector
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Knobbe, Martens, Olson & Bear
: STREET: 501 West Broadway
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-3505
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy Disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Win95
: SOFTWARE: word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/058,746
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Israelson, Ned A.
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER: GENSET.015A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10317 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: DOUBLE
: TOPOLOGY: CIRCULAR
: MOLECULE TYPE: synthetic DNA
: ORIGINAL SOURCE:
: ORGANISM: Cloning vector pGendel
: FEATURE:
: NAME/KEY: pGendel
: LOCATION: 1..10317
: FEATURE:
: NAME/KEY: Homology with X06404 comp1 (411..1668)
: LOCATION: 9..1266
: IDENTIFICATION METHOD: blastn against X06404
: FEATURE:
  
```

```
NAME/KEY: Kanamycin resistance gene CDS
LOCATION: 142..957
IDENTIFICATION METHOD: By homology to X06404
FEATURE:
NAME/KEY: Tn1000's right end
LOCATION: complement 1332..1371
IDENTIFICATION METHOD: blastn against X60200)
FEATURE:
NAME/KEY: Homology with U46017 (1-472)
LOCATION: 1423..1894
IDENTIFICATION METHOD: blastn against U46017
FEATURE:
NAME/KEY: single stranded DNA replication origin
LOCATION: 1423..1894
IDENTIFICATION METHOD: By homology to U46017
OTHER INFORMATION: mutation T -> C 1658
FEATURE:
NAME/KEY: Homology with U51113 (2382..6997)
LOCATION: 1896..6544
IDENTIFICATION METHOD: blastn against U51113
FEATURE:
NAME/KEY: OriS
LOCATION: 1972..2188
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: repELR
LOCATION: 2897..2918
OTHER INFORMATION: Described in seqid 16
FEATURE:
NAME/KEY: RepE
LOCATION: 2903..3034
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: T3
LOCATION: 3043..3059
OTHER INFORMATION: Described in seqid 17
FEATURE:
NAME/KEY: LRT3RA
LOCATION: complement 3045..3069
OTHER INFORMATION: Described in seqid 15
FEATURE:
NAME/KEY: IncC
LOCATION: 3070..3320
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: Insertion 33 bases 3038..3071
FEATURE:
NAME/KEY: Para
LOCATION: 3655..4821
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: mutation G -> A 3878
FEATURE:
NAME/KEY: ParB
LOCATION: 4821..5792
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: ParC
LOCATION: 5865..6382
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: Homology with J01688 (complement 175..819)
LOCATION: 6574..7218
IDENTIFICATION METHOD: blastn against J01688
OTHER INFORMATION: mutation A -> G 7096
FEATURE:
NAME/KEY: CDS streptomycin sensitivity gene
LOCATION: complement 6716..7090
IDENTIFICATION METHOD: By homology to J01688
OTHER INFORMATION: mutation A -> G 6728
OTHER INFORMATION: mutation G -> C 6821
OTHER INFORMATION: mutation C -> T 6866
OTHER INFORMATION: mutation T -> C 7013
OTHER INFORMATION: mutation T -> A 7058
FEATURE:

NAME/KEY: ipSLR
LOCATION: 7155..7174
OTHER INFORMATION: Described in seqid 12
FEATURE:
NAME/KEY: SP6
LOCATION: 7230..7248
OTHER INFORMATION: Described in seqid 13
FEATURE:
NAME/KEY: Tn1000's left end
LOCATION: 7252..7291
IDENTIFICATION METHOD: blast (X60200)
FEATURE:
NAME/KEY: Homology with X02730 (complement 37..1959)
LOCATION: 7305..9227
IDENTIFICATION METHOD: blastn against X02730
FEATURE:
NAME/KEY: CDS levansucrase gene
LOCATION: complement 7379..8800
IDENTIFICATION METHOD: By homology to X02730
OTHER INFORMATION: mutation T -> C 7466
OTHER INFORMATION: mutation A -> G 7739
OTHER INFORMATION: mutation T -> C (Asn -> Asp) 8347
OTHER INFORMATION: mutation T -> C 8600
OTHER INFORMATION: mutation G -> A (Ala -> Val) 8772
FEATURE:
NAME/KEY: SLR3
LOCATION: 8711..8731
OTHER INFORMATION: Described in seqid 14
FEATURE:
NAME/KEY: Homology with J01636 (complement 1158..1465)
LOCATION: 9298..9623
IDENTIFICATION METHOD: blastn against J01636
FEATURE:
NAME/KEY: CDS alpha peptide beta-galactosidase
LOCATION: complement 9276..9497
IDENTIFICATION METHOD: By homology to J01636
FEATURE:
NAME/KEY: primer HE1
LOCATION: complement 9465..9479
FEATURE:
NAME/KEY: primer HE2
LOCATION: 9461..9475
FEATURE:
NAME/KEY: primer lacRS2AVT
LOCATION: complement 9603..9630
FEATURE:
NAME/KEY: primer lacE2M1u
LOCATION: 9289..9314
FEATURE:
NAME/KEY: Homology with M77789 (1889..2576)
LOCATION: 9629..10315
IDENTIFICATION METHOD: blastn against M77789
FEATURE:
NAME/KEY: high copy-number double-stranded DNA replication origin
LOCATION: complement 9629..10315
IDENTIFICATION METHOD: By homology to M77789
OTHER INFORMATION: mutation C -> T 9603
OTHER INFORMATION: site ScaI 10029 - 10034
OTHER INFORMATION: site PmlI 10038 - 10043
OTHER INFORMATION: CLONING SITES 10031 - 10041
FEATURE:
NAME/KEY: oriLRd
LOCATION: 9856..9881
OTHER INFORMATION: Described in seqid 8
FEATURE:
NAME/KEY: OSI
LOCATION: 10009..10026
OTHER INFORMATION: Described in seqid 10
FEATURE:
NAME/KEY: ORI
LOCATION: complement 10046..10062
OTHER INFORMATION: Described in seqid 11
FEATURE:
```

NAME/KEY: OrlRr
 LOCATION: complement 10182..10202
 OTHER INFORMATION: Described in seqid 9
 US-09-058-746-1

Query Match 25.2%; Score 607.2; DB 3; Length 10317;
 Best Local Similarity 66.3%; Pred. No. 1.5e-158;
 Matches 993; Conservative 0; Mismatches 448; Indels 57; Gaps 6;

```

OY 358 ACGAGACATCAAAAGATGGCGGAAAGACAGCTGCCCTAACCTTGACAGCTCTCTTA 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8803 ACGAGAACATCAAAAGATTTGGCAAAACAGTAACAGTAATTAACCTTTACTACGCACTG 8744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 418 TTAGTAGGAGCGGAGACCGCATATTTTGGCGAGCAATGAACGCGGAGCTACAG 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8743 CTGGAGAGAGG---CGCACTCAACCGTTTGGGAAGAAGAACACAAAGCATATAG 8687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 478 GAAGACTATGTTTGGCCATATTAACGCGCTGACAGCTDAAAATTCACAGACACA 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8686 GAACATACGCGCATTTCCCATATTAACGCGCATGATGCAAAATCCCTGAACAGCAA 8627
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 538 AACAGTCCCATTTAAAGTGCCCTCATTCATGATGACAGCAATCAAAACATTGATCG 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8626 AAAATGAAAATATCAAGTTCCTGATTCGATTCGTCACAAATTAATATCTCTCT 8567
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 598 GCAAAAGGCTATGATAGTCAAGGCAACTTAATAGATTAGATTGGATAGCTGGCCA 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8566 GCAAAAGC-----CTGGAGCTTTGGGAGACGCTGGCCA 8534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 658 CTGCAAAACGCTGATGCTACTGCGCAATTAATCATGATATCATGCTCTCCCTTTA 717
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8533 TTACAAAAGCTGAGCGACCTGCGCAACTATACGCGCTACCGCATGCTCTTTCATTA 8474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 718 GCAGGTGACCCCAAAACAGTATGATCTACTCTCATTTATTTATTCAAAAGTGGGT 777
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8473 GCGGAGATCTCTAAAAATGCGGATGACATCGATTTTACATGTTTATCAAAAATGCGC 8414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 778 GATACATGATTTGACGTGAAAAAATGCTGAAAGATTTTGAAGATATGATTAATTT 837
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8413 GAAACTTTATTTGACGTGAAAAAAGCTGCGCCGCTTTAAAGACAGGACAAATTC 8354
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 838 GTTCCAATGATCCGTATCTTAATATATCAACACAGAGTGTCTGCTCTACTTTA 897
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8353 GATGAGATGATTTCTATCTTAATAAGACCAACACAAAGATGCTGAGTCCACCATTT 8294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 898 ACCAAGATGAGCAAGTCCGTTTATTTATACAGATTACTAGTAATCTGAAGATGCT 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8293 ACATCTGACGAAAAATCCGTTTATTTCTACACTGATTTCTCGGTAACATTA----- 8241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 958 GGAACGGGTGCTGTAACCAAAATCATTTCACTGCTCAAGTAACCTTATCCAGCGGAT 1017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8240 -----CGGCAAAACAAACACTGACAACTGACAAAGTAACGTATCAGCATCAGAC 8192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1018 GCAGCTACATTAAGTGAAGTGAATCTGATCTATCTAATATCTGTTTGAAGCGGAGAC 1077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8191 AGCTCT---TGAACATCAACGCTGTAGAGATTATTAATCAATCTTTGA---CGGTAC 8138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1078 GGTACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGCAAGTGAATTCAGGTGAT 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8137 GGA AAAAGCTATCAAAATATGTACAGAGTTTCAATGATGAAGCACTACAGCTCAGCGAC 8078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1138 AACCATCTTTAAGACCCCTCACTATGTTGAAGTAAGGCCATTAATCTGCTCTT 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8077 AACCATCTGAGAGATCTCTCACTAGTAGAAGTAAGGCCAACAAATCTACTATTT 8018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1198 GAAGGATTAAGTGAACACAGATGTTATCAAGCGCATCTTTCAATATTAAGCT 1257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8017 GAAGCAAACTGAACTGAAGATGCTACCAAGCGCAAGATCTTTTAAACCAACA 7958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1258 TACTATGCGGAGAGTGAAGTCTCTTCCAGATGAAGAAAAAATAAATCTGCTCAAGTCT 1317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7957 TACTATGCAAAAGACATCATTTCTTCCGTCAAGAAATCAAAACCTTGCAGAAAGCAT 7898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

OY 1318 AAAAAACAATTCCTCTTACGCAATGATGATTAAGCATTTGTAATGGCCGATGAC 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7897 AAAAAACGCGCTGATTAAGCAAGCGGCTCTCGTATGATTAAGCTTAAGCATGAT 7838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1378 TATACAGTGAAGTGTATGAACCATTAAGTCTGATCAACACAGTATGAGATGATC 1437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7837 TACACACTGA AAAAGATGATTAACCGCTGATTCATCAACAGTAAAGATTAAT 7778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1438 GAACGCGCATATTAATTAATGAATTAATTAATGATATCAAGATTAAGGA 1497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7777 GAACGCGGACGCTTTAAATGAACGCGCAAAATGATATCTTCACTGACTCCCGGGA 7718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1498 TC AAAATGACAGTGAATTAACACAAAGATTTATATGTAAGGCCCGGAGGC 1557
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7717 TCAAAATGACAGATTAAGCAGCATTAAGCTTAAGATTTATTCATGTTATGTTCT 7658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1558 GATCTCTTAATGCGCCACACACCCGTAATGAATGATGACTTTGATGAATGAT 1617
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7657 AATCTTTAACTGCGCCATACAAAGCCGCTGAACAAATGCGCTGTGTTAAATGAT 7598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1618 CTGACCCCTGCTGATCTCACACACTTACTCTGATTTGCGGATCCGACCCGGAAGT 1677
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7597 CTGATCTTAAGCATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 7538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1678 AATAATGTGATCTACAGTTAATGAAGATTAAGGCTTATCCAGAACATCACTCT 1737
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7537 AACATGCTGATTAAGCTATATGACAAAGAGATTTCTACGACACAAACATCA 7478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1738 CACCTGCGGACAGCTTGGGTTAATTTAAAGGCTGACATCTGGAGGAAAT 1797
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7477 ACGTTTGGCGGACCTTCTCTGACATCAAAAGCAAGAAATCTGTTGCAAGAC 7418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1798 AGTTCC---GGACAGACCAATTCACATGATTAATCTCCATTAAGAAATGTC 1852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7417 AGCATCTTGAACAAAGCATTAACAGTTAACAATTAACCAAGCAAAAGAAATGCC 7360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6
 US-09-438-142-1/c
 Sequence 1, Application US/09438142
 Patent No. 6258571

GENERAL INFORMATION:
 APPLICANT: Ilya Chumakov
 APPLICANT: Hiroaki Tanaka
 TITLE OF INVENTION: High Throughput DNA sequencing vector
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe, Martens, Olson & Bear
 STREET: 501 West Broadway
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-3505

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: win95
 SOFTWARE: word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/438.142

FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Israelson, Ned A.
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: GENSET. 015C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10317 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double

TOPOLOGY: circular
MOLECULE TYPE: synthetic DNA
ORIGINAL SOURCE:
ORGANISM: Cloning vector pGendel
FEATURE:
NAME/KEY: pGendel
LOCATION: 1..10317
FEATURE:
NAME/KEY: Homology with X06404 compl (411..1668)
LOCATION: 9..1266
IDENTIFICATION METHOD: blastn against X06404
FEATURE:
NAME/KEY: Kanamycin resistance gene CDS
LOCATION: 142..957
IDENTIFICATION METHOD: By homology to X06404
FEATURE:
NAME/KEY: Tn1000's right end
LOCATION: complement 1332..1371
IDENTIFICATION METHOD: blastn against X06200)
FEATURE:
NAME/KEY: Homology with U46017 (1-472)
LOCATION: 1423..1894
IDENTIFICATION METHOD: blastn against U46017
FEATURE:
NAME/KEY: single stranded DNA replication origin
LOCATION: 1423..1894
IDENTIFICATION METHOD: By homology to U46017
OTHER INFORMATION: mutation T -> C 1658
FEATURE:
NAME/KEY: Homology with U51113 (2382..6997)
LOCATION: 1896..6544
IDENTIFICATION METHOD: blastn against U51113
FEATURE:
NAME/KEY: OriS
LOCATION: 1972..2188
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: repELR
LOCATION: 2897..2918
OTHER INFORMATION: Described in seqid 16
FEATURE:
NAME/KEY: RepE
LOCATION: 2903..3034
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: T3
LOCATION: 3043..3059
OTHER INFORMATION: Described in seqid 17
FEATURE:
NAME/KEY: LRT3RA
LOCATION: complement 3045..3069
OTHER INFORMATION: Described in seqid 15
FEATURE:
NAME/KEY: IncC
LOCATION: 3070..3320
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: insertion 33 bases 3038..3071
FEATURE:
NAME/KEY: Para
LOCATION: 3635..4821
IDENTIFICATION METHOD: By homology to U51113
OTHER INFORMATION: mutation G -> A 3878
FEATURE:
NAME/KEY: ParB
LOCATION: 4821..5792
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: ParC
LOCATION: 5865..6382
IDENTIFICATION METHOD: By homology to U51113
FEATURE:
NAME/KEY: Homology with J01688 (complement 175..819)
LOCATION: 6574..7218

IDENTIFICATION METHOD: blastn against J01688
OTHER INFORMATION: mutation A -> G 7096
FEATURE:
NAME/KEY: CDS streptomycin sensitivity gene
LOCATION: complement 6716..7090
IDENTIFICATION METHOD: By homology to J01688
OTHER INFORMATION: mutation A -> G 6728
OTHER INFORMATION: mutation G -> C 6821
OTHER INFORMATION: mutation C -> T 6866
OTHER INFORMATION: mutation T -> C 7013
OTHER INFORMATION: mutation T -> A 7058
FEATURE:
NAME/KEY: fpeSLR
LOCATION: 7155..7174
OTHER INFORMATION: Described in seqid 12
FEATURE:
NAME/KEY: SP6
LOCATION: 7230..7248
OTHER INFORMATION: Described in seqid 13
FEATURE:
NAME/KEY: Tn1000's left end
LOCATION: 7252..7291
IDENTIFICATION METHOD: blast (X06200)
FEATURE:
NAME/KEY: Homology with X02730 (complement 37..1959)
LOCATION: 7305..9227
IDENTIFICATION METHOD: blastn against X02730
FEATURE:
NAME/KEY: CDS levansucrase gene
LOCATION: complement 7379..8800
IDENTIFICATION METHOD: By homology to X02730
OTHER INFORMATION: mutation T -> C 7466
OTHER INFORMATION: mutation A -> G 7739
OTHER INFORMATION: mutation T -> C (Asn -> Asp) 8347
OTHER INFORMATION: mutation T -> C 8600
OTHER INFORMATION: mutation G -> A (Ala -> Val) 8772
FEATURE:
NAME/KEY: SLR3
LOCATION: 8711..8731
OTHER INFORMATION: Described in seqid 14
FEATURE:
NAME/KEY: Homology with J01636 (complement 1158..1465)
LOCATION: 9298..9623
IDENTIFICATION METHOD: blastn against J01636
FEATURE:
NAME/KEY: CDS alpha peptide beta-galactosidase
LOCATION: complement 9276..9497
IDENTIFICATION METHOD: By homology to J01636
FEATURE:
NAME/KEY: primer HE1
LOCATION: complement 9465..9479
FEATURE:
NAME/KEY: primer HE2
LOCATION: 9461..9475
FEATURE:
NAME/KEY: primer lacI52Avr
LOCATION: complement 9603..9630
FEATURE:
NAME/KEY: primer lacE2Mlu
LOCATION: 9289..9314
FEATURE:
NAME/KEY: Homology with M77789 (1889..2576)
LOCATION: 9629..10315
IDENTIFICATION METHOD: blastn against M77789
FEATURE:
NAME/KEY: high copy-number double-stranded DNA replication origin
LOCATION: complement 9629..10315
IDENTIFICATION METHOD: By homology to M77789
OTHER INFORMATION: mutation C -> T 9803
OTHER INFORMATION: site ScaI 10029 - 10034
OTHER INFORMATION: site PmlI 10038 - 10043
OTHER INFORMATION: CLONING SITES 10031 - 10041
FEATURE:

```

NAME/KEY: oriLrd
LOCATION: 9856..9881
OTHER INFORMATION: Described in seqid 8
FEATURE:
NAME/KEY: OSI
LOCATION: 10009..10026
OTHER INFORMATION: Described in seqid 10
FEATURE:
NAME/KEY: ORI
LOCATION: complement 10046..10062
OTHER INFORMATION: Described in seqid 11
FEATURE:
NAME/KEY: oriLrd
LOCATION: complement 10182..10202
OTHER INFORMATION: Described in seqid 9
US-09-438-142-1

```

```

Query Match      25.2% Score 607.2; DB 4; Length 10317;
Best Local Similarity 66.3%; Pred. No. 1.5e-158;
Matches 993; Conservative 0; Mismatches 448; Indels 57; Gaps 6;

```

```

OY 358 ACGATGACCTCAAAAGATGCGGCAAAAGAGCGCTGCGTACCTTCAGGACCTCTATA 417
DB 8803 ACGATGACCTCAAAAGATGCGGCAAAAGAGCGCTGCGTACCTTCAGGACCTCTATA 417
OY 418 TTAGTAGAGCGGAGCGAGCGCGCATATTTTGGCGAGCAATGACAGCGGAGCTACAG 477
DB 8743 CTGGAGGAGG---CGCAACTCAACGCTTTGGGAGAAAGAACCAAGCAATATAG 8687
OY 478 GAGACTATGTTTGGCCATATTAACGCGCTGACATGCTAAATTCAGAGACAA 537
DB 8686 GAAACATAGCGCATTTCCCATATTTACGCGCATGATGCTGCAATCCCTGACAGCA 8627
OY 538 AACACTCCATTAATTAAGTGGCTCATTCATGATGATGATGATGATGATGATG 597
DB 8626 AAAAAATGAAAAATATCAAGTTCTGATGCTGCTCACAATTAATAATCTCTCT 8567
OY 598 GCAAAAGGCTATGATAGTACGAGCAACTTAATAGATTAGATTGAGATTAGTGCCA 657
DB 8566 GCAAAAGG---CTGGAGCTTTGGGAGACGCTGGGCA 8534
OY 658 CTGCAAAAGCGTATGATGCTGCGGCAATTAATATCATGATGATGATGATGATG 717
DB 8533 TTACAAAAGCGTATGATGCTGCGGCAATTAATATCATGATGATGATGATGATG 717
OY 718 GCAGGTGACCAAAAGAGTATGATGATGATGATGATGATGATGATGATGATG 777
DB 8473 GCGGAGATCCTAATAATGCGGATGATGATGATGATGATGATGATGATGATG 8414
OY 778 GATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
DB 8413 GAAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8354
OY 838 GTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 897
DB 8353 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8294
OY 898 ACCAAGATGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATG 957
DB 8293 ACATGATGCGGAAAAATCCGTTATGATGATGATGATGATGATGATGATGATG 8241
OY 958 GGAACCGGTGCTGATACCAATTCATTCATGCTCAAGTAACTTATCCAGCGGAT 1017
DB 8240 -----CGGCAAAACAACTGACACTGACAAAGTAAAGTATGACATCAGAC 8192
OY 1018 GCAGTACACTTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1077
DB 8191 AGCTCT---TTGAACATACAGGCTGTAGAGATTATTAATCATCTTTGA---CGGTGAC 8138
OY 1078 GTACAGTTTATCAAAATATTCAGCAATTTATGATGATGATGATGATGATGATGAT 1137
DB 8137 GGAACCAAGCTATCAAAATATTCAGCAATTTATGATGATGATGATGATGATGATG 8078

```

```

OY 1138 AACCATCTTTAGAGACCCCTACTATGTTGAGATAGAGCCATAATATCTTCTTT 1197
DB 8077 AACCATAGCGTGAAGATCTCTACTAGTGAAGATAGATAGATAGATAGATAGATAG 8018
OY 1198 GAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1257
DB 8017 GAAGCAAACTGAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATG 7958
OY 1258 TACTATGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1317
DB 7957 TACTATGCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7898
OY 1318 AAAAAAATGCTTCTTTAGGCAATGCTGATGATGATGATGATGATGATGATGATG 1377
DB 7897 AAAAAAGCAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7838
OY 1378 TATACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437
DB 7837 TACACACTGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 7778
OY 1438 GAAGCGGCAATATATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1497
DB 7777 GAAGCGGCAATATATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 7718
OY 1498 TCAAAATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557
DB 7717 TCAAAATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7658
OY 1558 GACTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1617
DB 7657 AATCTTTAACTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 7598
OY 1618 CTGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1677
DB 7597 CTGATCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7538
OY 1678 AATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1737
DB 7537 AACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7478
OY 1738 CACCTGCGGAGCAAGCTTGGGTTAATTAATTAAGGCTGATGATGATGATGATGATG 1797
DB 7477 ACGTTTGGCGGAGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATG 7418
OY 1798 AGTTCC---GGACAAAGCAATTCCTCAGATGATGATGATGATGATGATGATG 1852
DB 7417 AGCATCTTGAACAAAGCAATTAACGTTAAACAAATTAACCAAGCAAGCAAGCA 7360

```

```

RESULT 7
US-09-068-043-1
; Sequence 1, Application us/09068043
; Patent No. 6048694
;
; GENERAL INFORMATION:
; APPLICANT: MICHAEL GENE BRAMUCCI
; APPLICANT: VASANTHA NGARAJAN
; TITLE OF INVENTION: POSITIVE SELECTION
; TITLE OF INVENTION: VECTOR FOR BACILLUS SP.
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND
; ADDRESSEE: COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT OFFICE 97
; CURRENT APPLICATION DATA:

```


APPLICATION NUMBER: US/09/068,043
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/006,201
 FILING DATE: NOVEMBER 3, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: FELTHAM, S. NEIL
 REGISTRATION NUMBER: 36,506
 REFERENCE/DOCKET NUMBER: CR-9807
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-992-6460
 TELEFAX: 302-773-0164
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3305 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 US-09-068-043-1

Query Match 23.0%; Score 553.8; DB 3; Length 3305;
 Best Local Similarity 64.6%; Pred. No. 6,1e-144;
 Matches 932; Conservative 0; Mismatches 457; Indels 54; Gaps 5;

```

QY 357 GACGATGACTTCAAAAGATTGGCGAAAAAGCAGCTGCCCTACCTTCAGAGCTGCTAT 416
DB 1801 GACATGTGACATCAAAAAATTGTAAACACAGCAGTACGACTAGCTTTACGCACT 1860
QY 417 ATTACTAGAGCGAGCGACCGCATATTTTTCGAGCAAAATGACAGCGGGAGCTACAA 476
DB 1861 GCTACACAGA---GGAGCAACCTACAGCCTTCTGGAAAGAAATTAACCAAAAGCATACAA 1917
QY 477 GGAAGACTATGTTTGGCCATATTACACGCGCTGACATGCTAAAAATTCAGACAACA 536
DB 1918 AGAAACGTACGCGCTCTCATATTACACGCCCATGATATGTCAGATCCCTAAACAGCA 1977
QY 537 AAACGCTCTCAATTAAATGACCTCAATTCATGATGATGATGATGATGATGATGATG 596
DB 1978 GCAAAACGAAATATCCAAAGTGCCTCAATTCGATCAATCAAGATTAATAAATTTGAGTC 2037
QY 597 GGCAAAAGGATATGATAGTACAGCACTTAATAGATTAGATGATGATGATGATGATG 656
DB 2038 TGCAAAAG-----ACTGATGCTGTCGACAGCTGCGC 2070
QY 657 ACTGCAAAACGCTGATGCTACTGCGCAAAATTCATGATGATGATGATGATGATG 716
DB 2071 GCTGCAAAACGCTGACGAGACAGTACGAAATACAAAGGCTATCAGTTGTTGCTCT 2130
QY 717 AGCAGTGACCCAAAACAGTATGATGATGATGATGATGATGATGATGATGATGATG 776
DB 2131 TCGGGGAAGCCCGAAAGAGCGTGAACATCATCATGTTTATCAAAAGGTGCGG 2190
QY 777 TGATACATCATGATGACAGCTGGAATAATGATGATGATGATGATGATGATGATG 836
DB 2191 CGACAATCATGATGACAGCTGGAATAATGATGATGATGATGATGATGATGATG 2250
QY 837 TGTTCCAATGATGCTGATCTTAATATCAACACAGAGTGTGAGGTCTGCTACTTT 896
DB 2251 CGAGCCCAACGATCGATCTGGAAGATCAGACGCAAGAAAGTCCGGTTCGCAACCT 2310
QY 897 AACCAAGATGGCCAGTCCGTTTCTATACAGATTAATGATGATGATGATGATG 956
DB 2311 TACATCTGACGAAAAATCCGTTATCTACAGTACTATTCGGTAAACATTA----- 2364
QY 957 TGGAAACCGGCTGTGTAACCAATCTTCAACTGCTCAAGTAATCTTCCAGCCGGA 1016
DB 2365 -----CGCAAAACAAAGCTGACAAACAGCAGGTAATGTGTCAAAATCTGA 2412
QY 1017 TGCAGTACACTTAAAGTCAGTGAAGTATCTGATCATTAATCTGCTTTGATGCGGAGA 1076

```

```

DB 2413 T---GACACACTCAAAATCAACGAGTGAAGATCACAAAAGATTTTGA---CGGAGA 2466
QY 1077 CGGTACAGTTTATCAAAATATTCACCAATTTATGATGAAGCAGTGGATTTCAGTGA 1136
DB 2467 CGGAAAACATATACAAACGTTTCACAGTTTATGATGAAGCAATTTATCATTCGCGGA 2526
QY 1137 TACCATCTTTAAGACCCCTCAGTATGTAAGTAAGGCCATTAATATCTGTCTT 1196
DB 2527 CAACCATACGCTGAGAGACCCCTCAGTATGTAAGTAAGGCCATTAATATCTGTCTT 2586
QY 1197 TGAAGCGAATCTGGAACACAGATGTTATCAAGCGCATGCTTTCAATATTAAGC 1256
DB 2587 CGAAGCCAAACAGGGAACAGAAAGGATACCAAGCGGAATCTTTATTAATTAAGC 2646
QY 1257 TTACTATGGCGGAAGTACGCTTCTTCAGATGTAAGTAAGTAAGTAAGTAAGTCC 1316
DB 2647 GTACTACGGCGGCGACGACACTTCTTCTGTAAGTAAGTAAGTAAGTAAGTCC 2706
QY 1317 TAAAAAACAATGCTTTTATGCAATGATGATGATGATGATGATGATGATGATG 1376
DB 2707 TAAAAAAGCGGATGCTGATGATGATGATGATGATGATGATGATGATGATG 2766
QY 1377 CTATACAGTGAAGAGTGTATGAACACATTTAGTGCATCAACACAGATGATGAAGT 1436
DB 2767 TTACACATTTGAAGAAAGTATGATGATGATGATGATGATGATGATGATGATG 2826
QY 1437 GGAAGCGGCGCATATATTTAAATGATATTAATGATGATGATGATGATGATGATG 1496
DB 2827 CGAGCGCGCGAATGTTTCAAAATGATGATGATGATGATGATGATGATGATGATG 2886
QY 1497 ATCCAAATGACAGTATGATGATGATGATGATGATGATGATGATGATGATG 1556
DB 2887 TTCAAAATGACAGTATGATGATGATGATGATGATGATGATGATGATGATG 2946
QY 1557 GCATCTCTTAATGAGCCACACACACACACACACACACACACACACACACACAC 1616
DB 2947 AAACCTTTAAACCGCCCTTACAAAGCGCTGAAACAAAGCGCTTGTCTCAATGGG 3006
QY 1617 TCTTGACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1676
DB 3007 TCTTGATTCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3066
QY 1677 TAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1736
DB 3067 CAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3126
QY 1737 TCACCTGCGGGAACAAGCTTGGGTTAATTAAGGCTGTGACACATCTGAGAGAAAA 1796
DB 3127 AACATTTGGCCCAAGCTTCTTAATCAACATCAAAAGCAATAAATCATCGTTGCAAAA 3186
QY 1797 TAG 1799
DB 3187 CAG 3189

```

RESULT 8
 US-08-232-463-14/c
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFLINGER, F.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/232,463
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)83-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

Query Match 1.9%; Score 44.8; DB 1; Length 7218;
Best Local Similarity 8.8%; Pred. No. 0.02;
Matches 40; Conservative 211; Mismatches 203; Indels 0; Gaps 0;
QY 289 CTATCTTTTATATAGAACTATCAAGAGCTTCTTATCAATTCATACATATCCA 348
DB 1475 CTATCTATGCAAGTGTAAAGATAGAGATTTGGTACRRRRRRRRRRRRRRRR 1416
QY 349 AGGAGGAGAGATCACTCAAAAGATTGGGCAAAAAGCGCTGCGTAACCTCAGG 408
DB 1415 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1356
QY 409 ACTGCTATATAGTAGAGAGCGAGCGACCATATTTTGGCGCAAAATGAAACGCGG 468
DB 1355 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1296
QY 469 GACTTCAAGGAAGACTATGCTTTGCCATATTTACAGCGCTGACATCTAATAATTCCA 528
DB 1295 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1236
QY 529 GGACAAACAAACAGCTCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAC 588
DB 1235 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1176
QY 589 ATTGATTCGGCAAAAGGTATGATAGTCAGCAACTTAATAGATTAGATGATGATG 648
DB 1175 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1116
QY 649 AGCTGGCACTGCAGAAAGCTGATGCTACTGCGCAAAATTATCATGATACATCTGC 708
DB 1115 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1056
QY 709 TCCGCTTAGACAGTGACCAAAAACAGTATG 742
DB 1055 CCGGAGCTGCAAGCAAGCTCGGAATTATCTG 1022

RESULT 9
US-08-671-947-1
Sequence 1, Application US/08671947
Patent No. 5827515
GENERAL INFORMATION:

APPLICANT: Shahabi Reynoso, Mlira
APPLICANT: Yamamoto, Takashi
APPLICANT: Cooper, Nicole H.
APPLICANT: Kalmen, Sue S.
TITLE OF INVENTION: BACILLUS THURINGIENSIS SPOROLATION GENE
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: SANDOZ AGRO, INC.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,947
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Marcus-Wyner, Lynn
REGISTRATION NUMBER: 34,869
REFERENCE/DOCKET NUMBER: 133-0724
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3588
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1662 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 474..1427
OTHER INFORMATION: /codon_start= 474
US-08-671-947-1

Query Match 1.7%; Score 39.8; DB 1; Length 1662;
Best Local Similarity 67.5%; Pred. No. 0.24;
Matches 56; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1452 ATTTAAATGATATATAATGATATTCAGCGATTCAAGAGATCCAAATGACGAG 1511
DB 652 ATATTAAGATATATAAAGAAATATATCTGATTTATGTAATAAAAAGACAGAG 711
QY 1512 TGATGAATTAAAGCAAGATG 1534
DB 712 AGAAGGATTAAGTCAAGAGAG 734

RESULT 10
US-09-276-531-82/c
Sequence 82, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 807 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT12
CLONE: 1617155
US-09-276-531-82

Query Match 1.6%; Score 38; DB 4; Length 807;
Best Local Similarity 57.9%; Pred. No. 0.55;
Matches 62; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1853 ACTGGCAAGATCTTACCTTTTCCCAACCTTTTAAATAAAGGGTTTGTACAC 1912
Db 250 ATTCCCAATTTCACCTGCGCCGCGCTTTTAAANAAAGGTCGCGTAAC 191

Qy 1913 CTTAGAAGAAAGATGCTTGGCCGCGCAATTACCATGCC 1959
Db 190 CTTGGGGGAAANACCCTTGCGCGGTAAANCCCAAAACC 144

RESULT 11
US-09-134-001C-650
Sequence 650, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 650
LENGTH: 1434
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-650

Query Match 1.6%; Score 37.6; DB 4; Length 1434;

Best Local Similarity 51.8%; Pred. No. 0.93;
Matches 85; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 84 AGAAGACGCTTACGACGACATTAAGAGATTTCAAAGCTCTGGGATAGGCC 143
Db 732 AGCCATTATCATTAATTTCACACCGTCACCTTAATAAAGATCATTTGTAATAT 791

Qy 144 AAGTCACCTTGAATAAATTAAGCTCTGTCGGGCTTTTGTATTATTTAT 203
Db 792 GCGTGTGTTAAATAAATTAAGTGTTCACACTTTCAGCTTTCTTCATGCTGTGTAATAT 851

Qy 204 TTCAACTGCAAGTGTCCATCCCTATATCATTTAAGACGAA 247
Db 852 TACGATGTAAGTCTCGCTTACATCATTAATTACATTCAAA 895

RESULT 12
US-08-480-604A-9
Sequence 9, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 42A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7101 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7098
US-08-480-604A-9

Query Match 1.5%; Score 36.6; DB 1; Length 7101;
Best Local Similarity 60.6%; Pred. No. 3.7;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 622 AACTTAATAGATTGATGTAGTGGTGGCCACTGCAAAACGCTGATGCTACTCGG 681
DB 3478 AATTCATAAGCTTTTGGTAATGTAATCTGAGAAATGGAAGGTGTTCAAGTCACTACT 3537

OY 682 GCAATTTATCATGATATCATCATCGTCTCCGCTTTAGCA 720
DB 3538 GTAACGTGATGATATGATGATCACTTCTTTTCAGCACCATCA 3576

RESULT 13

US-08-405-496A-9
Sequence 9, Application US/08405496A
Patent No. 5919665

GENERAL INFORMATION:

APPLICANT: WILLIAMS, JAMES A.

TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM

TITLE OF INVENTION: NEUTROXIN

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/405,496A

FILING DATE: 16-MAR-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPND-01308

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 7101 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 1..7098
US-08-405-496A-9

Query Match 1.5%; Score 36.6; DB 2; Length 7101;
Best Local Similarity 60.6%; Pred. No. 3.7;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 622 AACTTAATAGATTGATGTAGTGGTGGCCACTGCAAAACGCTGATGCTACTCGG 681
DB 3478 AATTCATAAGCTTTTGGTAATGTAATCTGAGAAATGGAAGGTGTTCAAGTCACTACT 3537

OY 682 GCAATTTATCATGATATCATCATCGTCTCCGCTTTAGCA 720
DB 3538 GTAACGTGATGATATGATGATCACTTCTTTTCAGCACCATCA 3576

RESULT 14

US-08-915-136-9
Sequence 9, Application US/08915136
Patent No. 6290960

GENERAL INFORMATION:

APPLICANT: KINK, JOHN A.

APPLICANT: THALEY, BRUCE S.

APPLICANT: PADHYE, NISHA V.

APPLICANT: FTRCA, JOSEPH R.

APPLICANT: STAFFORD, DOUGLAS C.

TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND

TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,136

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/480,604

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/405,496

FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPND-01763

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 9:

PD 17-DEC-1997.
XX 10-JUN-1997; 97EP-0304032.
XX 10-JUN-1996; 96JP-0170630.
XX (HAYB) HAYASHIBARA SEIBUTSU.KAGAKU.
PI Chae H, Kubota M, Tsusaki K.
DR WPI: 1998-034976/04.
XX P-PSDB: AAW48306.
PT Bacillus beta-fructofuranosidase (s) or alcohol(s)
XX fructo:furanosylated saccharide(s)
XX Claim 8; Pages 23-25; 32pp: English.
CC This is a genomic DNA sequence which encodes
CC beta-fructofuranosidase, an enzyme that has an optimum temperature
CC of 50 deg. C at pH 6. The enzyme was produced by recombinant DNA
CC technology. It catalyses fructofuranosyl transfer reactions from a
CC fructofuranosyl donor to a fructofuranosyl acceptor where the donor is
CC sucrose, raffinose or erlose and the acceptor is selected from alcohols,
CC sugar alcohols and saccharides having no beta-fructofuranosidic linkages,
CC especially where the reaction product is xylosyl fructoside, erlose,
CC isomaltosyl fructoside, lactosucrose or fructosyltrehalose.
CC Such reaction products have a satisfactory taste and sweetness, a
CC moderate viscosity and humectancy, an effective anticariogenic activity,
CC growth promoting activity for Bifidobacteria, mineral-absorption
CC promoting activity. The products can be used to improve the tastes and
CC textures of food products, cosmetics and pharmaceuticals and are useful
CC as sweeteners in the food and pharmaceutical industries.
SQ Sequence 2408 BP; 777 A; 492 C; 496 G; 643 T; 0 other:
Query Match 100.0%; Score 2408; DB 19; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGGGAAATACTATTCATTCATTCAGGAGTCCCGAGTGGTGTAAGAGAAGTTCG 60
DB 1 CGGGGAAATACTATTCATTCATTCAGGAGTCCCGAGTGGTGTAAGAGAAGTTCG 60
QY 61 GACTCCCATGCGAGCTGTGCGTAAGAAAACAGCTTACTGAGCAATTCAGAGAAT 120
DB 61 GACTCCCATGCGAGCTGTGCGTAAGAAAACAGCTTACTGAGCAATTCAGAGAAT 120
QY 121 TTCAAAGTCTGGATTAAGGCCAGTCCACTTGAATAAATTCAGCTCTGTGTG 180
DB 121 TTCAAAGTCTGGATTAAGGCCAGTCCACTTGAATAAATTCAGCTCTGTGTG 180
QY 181 GGGGGTTTTTGTATTTATTTCACTGCAAGTGGTCCATCCCTATTCATTTTA 240
DB 181 GGGGGTTTTTGTATTTATTTCACTGCAAGTGGTCCATCCCTATTCATTTTA 240
QY 241 GAGCAATTTCTATCATTCATCCATCCCAATTAAGTCTCTCTTACTTTTAA 300
DB 241 GAGCAATTTCTATCATTCATCCATCCCAATTAAGTCTCTCTTACTTTTAA 300
QY 301 TTAATTAAGAACTATCAAGAGCTTTCTTATCAATTCATCATATTCAGAGAGAG 360
DB 301 TTAATTAAGAACTATCAAGAGCTTTCTTATCAATTCATCATATTCAGAGAGAG 360
QY 361 ATGAATCTCAAAAGATTTGGCGAAAAAGAGCTGGCGTAACCTTGAGGCTCTATTA 420
DB 361 ATGAATCTCAAAAGATTTGGCGAAAAAGAGCTGGCGTAACCTTGAGGCTCTATTA 420
QY 421 TTAGAGAGCGAGGAGCCGATATTTTGGCAGCAAAATGAACAGGGGAGTACAGAG 480
DB 421 TTAGAGAGCGAGGAGCCGATATTTTGGCAGCAAAATGAACAGGGGAGTACAGAG 480
QY 481 GACTATGTTTTGCCCATATTTACACGCGCTGACATGCTAAAAATTCAGAGCAACAA 540
DB 481 GACTATGTTTTGCCCATATTTACACGCGCTGACATGCTAAAAATTCAGAGCAACAA 540

DB 481 GACTATGTTTTGCCCATATTTACACGCGCTGACATGCTAAAAATTCAGAGCAACAA 540
QY 541 AGTCTCAATTTAAAGTCCCTCATTCATGATCAGCAATCAAAAACATGATTCGCA 600
DB 541 AGTCTCAATTTAAAGTCCCTCATTCATGATCAGCAATCAAAAACATGATTCGCA 600
QY 601 AAAGGTATGATTAAGTCAAGCACTTATATGATTTAGATGATGATGATGATGATG 660
DB 601 AAAGGTATGATTAAGTCAAGCACTTATATGATTTAGATGATGATGATGATGATG 660
QY 661 CAAAACGCTGATGCTGACGCAAAATTCATGATGATGATGATGATGATGATGATG 720
DB 661 CAAAACGCTGATGCTGACGCAAAATTCATGATGATGATGATGATGATGATGATG 720
QY 721 GGTGACCCCAAAAACAGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 GGTGACCCCAAAAACAGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 ACATGATTTGACGCTGCAAAAATGCTGAGAGATTTGAGATTTGAGATTTGAGAT 840
DB 781 ACATGATTTGACGCTGCAAAAATGCTGAGAGATTTGAGATTTGAGATTTGAGAT 840
QY 841 CCAATGATCCGATATCTTAAATATCAAAACAGAGAGTGTGCTGCTACTTTAAC 900
DB 841 CCAATGATCCGATATCTTAAATATCAAAACAGAGAGTGTGCTGCTACTTTAAC 900
QY 901 AAAGTGGCCAAAGTCCGTTTATCTATACATTTACTAGTATCTCAGAAATGATG 960
DB 901 AAAGTGGCCAAAGTCCGTTTATCTATACATTTACTAGTATCTCAGAAATGATG 960
QY 961 ACCGCTGTGTAACCAATATTTCACTGCTCAAGTAACTTATCCACCGGATGCA 1020
DB 961 ACCGCTGTGTAACCAATATTTCACTGCTCAAGTAACTTATCCACCGGATGCA 1020
QY 1021 GCTACATTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 GCTACATTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ACAGTTTCAAAAATTCATTCAGCAATTTATCGATGAGCAAGTGTGATGATGAT 1140
DB 1081 ACAGTTTCAAAAATTCATTCAGCAATTTATCGATGAGCAAGTGTGATGATGAT 1140
QY 1141 CATACTTTAAGAGACCTCCTATGTTGAAGTAAAGGCCATTAATATCTTGTGAA 1200
DB 1141 CATACTTTAAGAGACCTCCTATGTTGAAGTAAAGGCCATTAATATCTTGTGAA 1200
QY 1201 GCGAATCTGGAACACAGATGTTATCAAGGCGATGCTTTTCAATTAATGAAGCT 1260
DB 1201 GCGAATCTGGAACACAGATGTTATCAAGGCGATGCTTTTCAATTAATGAAGCT 1260
QY 1261 TATGCGGAGAGTACGCTCTTCTCCAGATGAAAAAATTAAGTCTTCAAAAGTCT 1320
DB 1261 TATGCGGAGAGTACGCTCTTCTCCAGATGAAAAAATTAAGTCTTCAAAAGTCT 1320
QY 1321 AAACAATTTGCTCTTTAGCAATGTTGATGATGATGATGATGATGATGATGATG 1380
DB 1321 AAACAATTTGCTCTTTAGCAATGTTGATGATGATGATGATGATGATGATGATG 1380
QY 1381 ACAGTGAAGAGTGTATGAACCATTTAGTCCGATTAACACAGTGAAGATGAGTGA 1440
DB 1381 ACAGTGAAGAGTGTATGAACCATTTAGTCCGATTAACACAGTGAAGATGAGTGA 1440
QY 1441 CCGCCCAATATTTAAATGAATTAATGAATGATGATGATGATGATGATGATGATG 1500
DB 1441 CCGCCCAATATTTAAATGAATTAATGAATGATGATGATGATGATGATGATGATG 1500
QY 1501 AAATGACAGTGAAGATTAACGCAAAAGATTTATGATGATGATGATGATGATGATG 1560
DB 1501 AAATGACAGTGAAGATTAACGCAAAAGATTTATGATGATGATGATGATGATGATG 1560
QY 1561 TCTTTAATGCGCCACACACCGATTAATGAATGATGATGATGATGATGATGATGAT 1620
DB 1561 TCTTTAATGCGCCACACACCGATTAATGAATGATGATGATGATGATGATGATGAT 1620


```

Db 1561 TCCTTAAATGGCCACACACACCCGATTAATGAACTGACCTGTATGTAACATGAACTCT 1620
OY 1621 GACCTCGTGAATCTCACACACACTTACTCTCATTTGGGATGCCACCTGAAGATAT 1680
Db 1621 GACCTCGTGAATCTCACACACACTTACTCTCATTTGGGATGCCACCTGAAGATAT 1680
OY 1681 AATGTGTACTACACAGTTATATGAGAAATAGAGGCTTCTATCCAGAACATCTCTCAC 1740
Db 1681 AATGTGTACTACACAGTTATATGAGAAATAGAGGCTTCTATCCAGAACATCTCTCAC 1740
OY 1741 CTGGGGGACAGCTGGGGGTTAATATTAAGGCTGTACACATCTGGAGAGAAATAGT 1800
Db 1741 CTGGGGGACAGCTGGGGGTTAATATTAAGGCTGTACACATCTGGAGAGAAATAGT 1800
OY 1801 TCCGACACAGACATATCCATAGCATATCTCCCAATAAAAAGAAATGTCTACTGGCAA 1860
Db 1801 TCCGACACAGACATATCCATAGCATATCTCCCAATAAAAAGAAATGTCTACTGGCAA 1860
OY 1861 GATCTTACCTTTTCCCAACCTTTTAAAAATAAAGGTTTGTACACACCTTTAGAA 1920
Db 1861 GATCTTACCTTTTCCCAACCTTTTAAAAATAAAGGTTTGTACACACCTTTAGAA 1920
OY 1921 GAAAAAAGATTCCTGGGCGCCGCAATTAACCATTTGCGGAGAGCTAACCAATAAAA 1980
Db 1921 GAAAAAAGATTCCTGGGCGCCGCAATTAACCATTTGCGGAGAGCTAACCAATAAAA 1980
OY 1981 AGAAATTTTCCCTTTTAAAGGGGAGTCTTTTCTATCTGGGGTGGGATTTGTTG 2040
Db 1981 AGAAATTTTCCCTTTTAAAGGGGAGTCTTTTCTATCTGGGGTGGGATTTGTTG 2040
OY 2041 GCCCCACAGGGGACCTTTATTAATTAATAGCATGAATCTGGCGATTTTGGCTGGCTA 2100
Db 2041 GCCCCACAGGGGACCTTTATTAATTAATAGCATGAATCTGGCGATTTTGGCTGGCTA 2100
OY 2101 CTTATTAACAAAGCCGCCCAAAATATTAACCAAGCAATCTTAAGATAGCCGCT 2160
Db 2101 CTTATTAACAAAGCCGCCCAAAATATTAACCAAGCAATCTTAAGATAGCCGCT 2160
OY 2161 ACAGTTTACCAACACCTCTGAGAGGATGATATCGCCCTCAATACACTTTTCTGTTG 2220
Db 2161 ACAGTTTACCAACACCTCTGAGAGGATGATATCGCCCTCAATACACTTTTCTGTTG 2220
OY 2221 CCGACAATGAAAAATGATCCGACAGAGGCCATTTTGGGGGAAGTATCATTA 2280
Db 2221 CCGACAATGAAAAATGATCCGACAGAGGCCATTTTGGGGGAAGTATCATTA 2280
OY 2281 TTACCTTACAAACAAGACTATCCAGATGAAATGTAACGGAATGCGGCATGCAACGTC 2340
Db 2281 TTACCTTACAAACAAGACTATCCAGATGAAATGTAACGGAATGCGGCATGCAACGTC 2340
OY 2341 CGAAGATTTATTTGATGAGAGGAGGATGATCCATCCGAGATATACCAATAAAA 2400
Db 2341 CGAAGATTTATTTGATGAGAGGAGGATGATCCATCCGAGATATACCAATAAAA 2400
OY 2401 TGGTGATC 2408
Db 2401 TGGTGATC 2408

```

RESULT 2
ABK75353 standard; DNA; 1449 BP.

ABK75353;

13-AUG-2002 (first entry)

Bacillus licheniformis genomic sequence tag (GST) #2644.

Differential gene expression; genomic sequenced tag; GST; altered culture condition; environmental stress;

physiological provocation; ds.

OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
XX 27-MAR-2001; 2001US-279526P.
PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
PI Berka R. Clausen IG;
XX WPI; 2002-416684/44.
XX
DR
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array

Claim 4; SEQ ID NO 2644; 200pp; English.

The invention describes a method of monitoring differential expression of genes in a first Bacillus cell relative to expression of the genes in other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which Bacillus cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPRO at fltp.wipro.int/pub/published_pcl_sequences.

Sequence 1449 BP; 492 A; 341 C; 313 G; 303 T; 0 other;

Query Match 26.0%; Score 627; DB 24; Length 1449;
Best Local Similarity 66.5%; Pred. No. 5; Se-167;
Matches 957; Conservative 0; Mismatches 455; Indels 27; Gaps 3;

```

OY 361 ATGAACCTCAAAAGAATGCGCAAAAGACAGCTGCCGTACCTTGACGCTCTATATTA 420
Db 1 ATGAACATCAAAAACATATGCTTAATAAAGCGTATAGCCTTAACCGT---TGTCTGGCAGCT 57
OY 421 GTAGAGCGGAGCGAGCCGATATTTTGGCGACGAATGAAACAGCGGAGACTAAGGAA 480
Db 58 CTGGCCGGAGGTGCGCCGCAAACTTTGCAAAAGAACGCA-----GGATTAACAAGAA 111
OY 481 GACTATGTTTGGCCATATTTACAGCGCTGACATGCTAATAAATTCGAGCAACAAAC 540
Db 112 AGCTACGGAATTTCTCATATACAAAGACATGCTGAATAATTCGAGAGCAAAAG 171
OY 541 AGTCTCAATTAAGTGCTCAATTTCAATGATGACATGCAATCAAAACATGATTGGCA 600
Db 172 AGCGACAATTTAAAGTTCCTCAATTCGATCCGAAACAAATCAAAACATCCCTTCTCA 231
OY 601 AAGGCTATGATAGTCAAGCAACTTAATAGATTGATGATGAGATAGCTGGCCACTG 660

```

```

Db      232  AAGGGTAAACAAAATGAGAGCTGATTCGATTTAGAGTATGGAGACAGTCCGCTG 291
OY      661  CAAAGCGTGATGCTGCGGCAATATATCATGATATACATGCTCCGCTTTACA 720
Db      232  CAAATGCGGAGGAGGCTTGCTACATACCGAGCTCAAACTTGTTCGCGCTGGG 351
OY      721  GGTGACCCCAAAAACAGTATGATCTCCACTTCTTATTTATCAAAAAGTCGGTAT 780
Db      352  GGGCATTCGAAAAGAGCTGATGACACATCCATCTTTTGTCTATCAAAAAGAGCGGA 411
OY      781  ACATGATTTGACAGCTGGAATAATCTGAGAGATTTTGAAGATATGATTAATTTGT 840
Db      412  ACTTCTATCGACAGCTGGAATAAACCCGCGAGAGTGTAAAGACAGGCAAAATTTGT 471
OY      841  CCAATGATCCGTATCTTAATATGAAACACAGAGTGCTGAGTCTGCTCTTAAAC 900
Db      472  CCAGCGATCCGTACCTCAAAACATTAACACAGATGCTGAGTCTGCTGCA 531
OY      901  AAGATGCGCAAGTCCGTTTATTTATACAGATTTACTAGTAAATCTGAAGATGCTGA 960
Db      532  AAGAGCGGAAAAGTCCGACTGTTTACAGCTTTTCGCGCAGCAATA----- 581
OY      961  ACCGTGCTGTGTAACCAATCTTTCAACTGCTCAAGTAACTTATCCAGCGGATGCA 1020
Db      582  -----CGGCAACAGACAGCTGACAAACAGCTCAGTCAATTTCTCAGCGGATTCG 633
OY      1021  GCTACACTTAAAGTGTAGTATGATATCATTAATCTGTTTGTGAGCGGAGAGCGGT 1080
Db      634  GACACGCTCAAAATTTGACGCTGTACAAAGATCATTAATCGCTTTGACGCGCGCAGCG 693
OY      1081  ACAGTTTATCAAAATATTTAGCAATTTATCGATGAAGCGCAAGTGAATTCAGTGATAAC 1140
Db      694  ACGGTATACCAAAAGCTTCAGCAATTCATTGCAAGAAACTACAGCTCCGCGCACAC 753
OY      1141  CATACTTAAAGAGACCTTCACATTTGAAAGTAAGCGCCAAATATCTTCTTGA 1200
Db      754  CATACGATGAGAGAGCCGCAATTTATGTGAAGACCGCGCAATAATATCTGATTTGAA 813
OY      1201  GCGAATCTGGAACAACAGATGTTATCATAGCGGATCATCTTCAATTAATAACTTAC 1260
Db      814  GGCATACGGGAGCAAAAACCGGCTACCAAGAGAAAGACTCCCTTATTCACAGAGCTTAC 873
OY      1261  TATGCGGGAAGTGAAGTCTTCTCCAGATGAAAAAATACTCTTCAAGTCTTAAA 1320
Db      874  TACGGGGGACCAAGAAAGTCTTTTAAAGAAAGAACACAGACTCTGCAAGAGTCCGAA 933
OY      1321  AAACAATTTGCTCTTAGCAATGCTCATTAGCAATTTGATTTGGCCATACAT 1380
Db      934  AAAAAGAGCGCTCGCTGCTAAGCGCGCTCTCGGAATTCATCGAATTAATTAACGATTAT 993
OY      1381  ACAGTGAAGAAGTGTATGAACCATTTAGTCGATCAAAACAGAGTGAAGTGAAGTGA 1440
Db      994  ACACGAAAAAAGTATGATAGAGCTTTGATCGCTCCATACGGGTACAGATTAATTCGA 1053
OY      1441  CGCGCCAAATATTTAAATGAATATTAATGATATCTATTCACGATTCAGAGAGATTC 1500
Db      1054  CGGGCAACCTCTTCAAAATGAATGAATAATGATCTGTTACAGATTCAGAGAGATCA 1113
OY      1501  AAAATGAGAGTGAATGAATTAACACAAAGATGTTTATGCTAGGGCCCGGAGCGCAC 1560
Db      1114  AAAATGACAATTTGAGCGGATCGTTCAAAAGACATTTATATGCTGGGTATGATACAGT 1173
OY      1561  TCCTTAATGCGCCACACACCCGATTAATGAATGAATGAGTGTATTAACATGAATCTT 1620
Db      1174  TCATTAACCGGACCATTTAAGCTTTAAACAATCCGAGCTTTGTTTGCATATGAGACAG 1233
OY      1621  GACCTGCTGATCTCACACACACTTACTCATTTGCGGTATCCGACCTGAGAGTAT 1680
Db      1234  GATTACATGACATCAGCTTTACTTATTCACACTTTGCGGTACCGCAGAAAAAGCGCAC 1293
OY      1681  AATGTGATCTCACAAGTATATGAGATAGAGCTTCTATCCGAACATCATCTCTAC 1740
Db      1294  GAAGTGTATTTACAAAGCTACATCAAAACAGAGGATTTGAAAGACATCAGCCACG 1353

```

```

OY      1741  CTGCGGACAAGCTTGCGGTTAATATTAAGGCTGTACACATCGAGGAGAAAAATAG 1799
Db      1354  TTTGCACCAAGCTTTTCTCTGACATCAAAAGATCAAAACATCCGTTGTCAAAAACAG 1412

RESULT 3.
AA235988/C
ID      AA235988 standard: DNA: 10317 BP.
XX
AC      AA235988;
XX
DE      09-FEB-2000 (first entry)
XX
Vector pGendell sequence SEQ ID NO:1.
XX
KW      DNA sequencing vector; nested deletion; transposition; contig;
KW      truncated lacZ; origin of replication; pGendell; mapping; ss.
OS      Synthetic.
PN      WO9953044-A2.
XX
PD      21-OCT-1999.
XX
PF      09-APR-1999; 99WO-IB00690.
XX
PR      10-APR-1998; 98US-0058746.
XX
PA      (GEST ) GENSET.
XX
PI      Chumakov I, Tanaka H;
XX
DR      WPI: 2000-023040/02.
XX
PT      New DNA sequencing vectors, used for sequencing large regions of DNA
XX
PS      and mapping the location of markers
XX
XX
Claim 43; Page 93-96; 102pp; English.
XX
CC      The present invention describes a vector comprising the following:
CC      (a) a high copy number origin of replication (Ori) having at least one
CC      cloning site, with at least one cloning site being positioned in the
CC      high copy number Ori such that the ability of the high copy number Ori
CC      to direct replication is not disrupted when no insert has been cloned
CC      into the cloning site and is disrupted when an insert is cloned into
CC      the cloning site; and (b) a low copy number Ori. Also described is a
CC      truncated LacZ gene to distinguish cells with the truncated LacZ
CC      gene at high copy number from cells carrying it at a low copy number,
CC      where the former are dark blue when grown on medium containing Xgal and
CC      IPTG and the latter are light blue when grown on medium containing Xgal
CC      and IPTG. Products and methods from the present invention can be used
CC      for the manipulation of DNA. They can be used for sequencing large
CC      regions of DNA and mapping the locations of markers within large regions
CC      of DNA. The present sequence represent a specifically claimed vector
CC      sequence from the present invention.
XX
XX
SQ      Sequence 10317 BP; 2567 A; 2296 C; 2469 G; 2985 T; 0 other;

Query Match      25.2%; Score 607.2; DB 21; Length 10317;
Best Local Similarity 66.3%; Pred. No. 6.2e-161;
Matches 993; Conservative 0; Mismatches 448; Indels 57; Gaps 6;

OY      358  ACGATGACCTTCAAAAGATTTGGCGAAAAACAGCTGCGGTAACCTTACAGACGCTATA 417
Db      8803  ACGATGACATCAAAAAGTTTGCAAAACAAATACAGATTAACCTTTATACACGCACTG 8744
OY      418  TTAGTAGAGCGGAGCGACCCGATATTTTTCGCGAGCAAAATGAACAGCGGAGTACAG 477
Db      8743  CTGCGAGGAGG---CGCACTCAAGCGTTTTCGGAAGAAAGAAACCAAAAGCCATATAG 8687
OY      478  GAAGCATATGTTTGTCCCATATTACAGCGCTGACATGCTAAAATTCAGAGACAACAA 537

```

```

Db      8686 GAACATACGGCATTCCCAATATACACGCCATGATATGTCGCAAAATCCCTGAAACAGCA 8627
Oy      538 AACAGTCCCTCAATTTAAAGTCGCTCAATTCATGATGATGAGCAATCAAAACATTTATCG 597
Db      8626 AAAAATGAAAAATATCAAGTTCCTGATTCGATTCGATCCACATATAAAAATATCTCTCT 8567
Oy      558 GCAAAAGGGTATGATTAAGTACAGGCAACTTAATAGATTAGATGATGGAATAGCTGGCCA 657
Db      8566 GCAAAAG-----CTGAGCGTTTGGGACAGCTGGCCA 8534
Oy      658 CTGCAAAAGCGTGANGTACTGCGGCAATTAATCATGATATCATGCTGCTCCGCTTTA 717
Db      8533 TTCAAAAGCGTGANGTACTGCGGCAATTAATCATGATATCATGCTGCTCCGCTTTA 8474
Oy      718 GCAGGTGACCCCAAAACAGTATGATATCTCCATCTTATTTATTTATCAAAAAGTCCGT 777
Db      8473 GCGGAGATTCCTAAATAATGCGGATACACATCGATTTTACATCTTCTATCCAAAAGTCCGC 8414
Oy      778 GATCATCGATTGACAGCTGGAATAATGCTGGAAGATTAATTTGAAGATATGATTAATTT 837
Db      8413 GAAACTTCTATTTAGACGTGGAATAATGCTGGAAGATTAATTTGAAGATATGATTAATTT 8354
Oy      838 GTTCCAAATGATCCCTATCTTAATATATCAACACAGAGAGTGTGCTGCTACTTTA 897
Db      8353 GATGCAATGATTTCTATCTTAAAGACCAACACAAAGATGCTGCTGCTGCTGCTGCTGCT 8294
Oy      898 ACCAAGATGCGCAAGTCCGTTTATTTATACAGATTAATCTGATGATTAATCTGATGAT 957
Db      8293 ACATCTGACGGAATAATCCGTTTATTTATCTGATGATTAATCTGATGATTAATCTGATGAT 8241
Oy      958 GGAACCGGTGCTGTATCAAAATCTTTCAACTGCTGCAAGTAACTTATCCGACGCGAT 1017
Db      8240 -----CGGCAAAACAACTGACAACTGACAACTGACAACTGACAACTGACAACTGAC 8192
Oy      1018 GCAGTACACTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Db      8191 AGCTCT---TTGACATCAACGGTGTAGAGATTAATTAATCAATCTTTGA---CGGTGAC 8138
Oy      1078 GGTACAGTTTATCAAAATATTTACAGCAATTTATGATGATGATGATGATGATGATGATGAT 1137
Db      8137 GGAATAAGGTATCAAAATATTTACAGCAATTTATGATGATGATGATGATGATGATGATGAT 8078
Oy      1138 AACCATCTTTAAGAGACCTCTACTATGTTGAAGTAAAGGCGCATTAATATCTTCTTT 1197
Db      8077 AACCATCTTTAAGAGACCTCTACTATGTTGAAGTAAAGGCGCATTAATATCTTCTTT 8018
Oy      1198 GAAGCATCTGGAACAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
Db      8017 GAAGCATCTGGAACAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7958
Oy      1258 TACTATGCGGGAAGTGAAGTCTTCTTCCAGATGAAAAAATAAATCTGCTTCAAGTCT 1317
Db      7957 TACTATGCGGGAAGTGAAGTCTTCTTCCAGATGAAAAAATAAATCTGCTTCAAGTCT 7898
Oy      1318 AAAAAACAATGCTTCTTTAGCGAATGTCATTTAGCGATTTGAAATTTGCGCGATGAC 1377
Db      7897 AAAAAACAATGCTTCTTTAGCGAATGTCATTTAGCGATTTGAAATTTGCGCGATGAC 7838
Oy      1378 TATACAGTAAAGTGTATGAAACCATTTAGTGCATTAACACAGTATGACAGTATGATGAT 1437
Db      7837 TATACAGTAAAGTGTATGAAACCATTTAGTGCATTTAGTGCATTTAGTGCATTTAGTGCAT 7778
Oy      1438 GAAGCGCGCAATATATTTAAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1497
Db      7777 GAAGCGCGCAATATATTTAAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 7718
Oy      1498 TCCAAAATGACAGTATGATTAACGACAAAGATGTTTATGATGATGATGATGATGATGATGAT 1557
Db      7717 TCCAAAATGACAGTATGATTAACGACAAAGATGTTTATGATGATGATGATGATGATGATGAT 7658
Oy      1558 GATCCTTAATTTGCGCCACACACCGCATTAATGATGATGATGATGATGATGATGATGATGATGAT 1617
Db      7657 AATCTTTAACTGCGCATACAAAGCGCTGGAACAAACTGCGCTTGTGTTAAATGAT 7598

```

```

Oy      1618 CTTGACCGTGTGATCTCACACACACTTACTCTCATTTGCGGTATCCCGCACCTGAAAGT 1677
Db      7597 CTTGATCTCAAGATGATTAACCTTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 7538
Oy      1678 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737
Db      7537 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7478
Oy      1738 CACTGCGGGAACACTTGGGTTTAAATTAATTAAGGCTGACACATCTGAGAGAAAT 1797
Db      7477 ACGTTTGGCGGAGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7418
Oy      1798 ACTTCC---GCACAAGACAAATTTCCATGATGATGATGATGATGATGATGATGATGATGAT 1852
Db      7417 AGCATCTTGAACAAAGACAAATTAACATGATTAACAAATTAACAAATTAACAAATTAAC 7360

```

```

RESULT 4
AA257329
ID AA257329 standard; DNA; 1668 BP.
XX
AC AA257329;
XX
DT 03-APR-2000 (first entry)
XX
DE SacB and cpy chimeric gene #1.
XX
KW Levanisucrase; SacB; cpy; salt tolerance; vacuole guide peptide;
KW carboxypeptidase A; chimeric gene; transgenic plant; yeast;
KW Bacillus subtilis 168; Saccharomyces cerevisiae X8;
KW drought resistance; ds.
XX
OS Chimeric - Saccharomyces cerevisiae.
OS Chimeric - Bacillus subtilis.
XX
PN CN1231337-A.
XX
PD 13-OCT-1999.
XX
PF 08-APR-1998; 98CN-0101336.
XX
PR 08-APR-1998; 98CN-0101336.
XX
PA (GENE-) INST GENETICS CHINESE ACAD SCI.
XX
PI Chen S, Zhang H, Dong W;
XX
DR WPI, 2000-087902/08.
XX
PT Method for improving plant salt resistance using transgenic technology
XX
PS Claim 6; Page 2; 12pp; Chinese.
XX
CC A method has been developed for raising salt tolerance of plants using
CC transgenic technology. The method comprises: cloning levanisucrase gene
CC SacB and vacuole guide peptide (cpy) sequence of carboxypeptidase A
CC from Bacillus and DNA of yeast, using the two kinds of gene to create a
CC chimeric gene, and using the chimeric gene to structure plant expression
CC carrier, using the obtained plant expression carrier to transform a
CC plant and screen for resistant seedlings. The method can obtain
CC drought-resistant salt-tolerance plants, and the polymerase chain
CC reaction (PCR) and Northern analysis of transgenic plants shows that
CC the exogenous gene is integrated in the salt-tolerance transgenic plant.
CC The present sequence represents a SacB and cpy chimeric gene from the
CC present invention.
XX
SQ Sequence 1668 BP; 578 A; 372 C; 325 G; 393 T; 0 other;
XX
Query Match 24.0%; Score 577.8; DB 21; Length 1668;
Best Local Similarity 66.4%; Pred. No. 5.2e-153;
Matches 916; Conservative 0; Mismatches 412; Indels 51; Gaps 4;

```

```

OY 421 GTAGAGCGGACGAGCCGATATTTTTCGCGACAAATGAACGCGGGGACTACAGGAA 480
Db 304 GCAAGTGAAGAACTACAGCTTCGTGCAATAGAAAGAACCAAAAGCCATATATAGAA 363
OY 481 GACATGTGTTTGGCCATATTACAGCGGCTGACATGCTAAATAATTCGAGACAAAC 540
Db 364 ACATACGGCATTTCCCATATTACAGCGCATGATATGTCGCAAAATCCGAAACGCAAAA 423
OY 541 AGTCCTCAATTTAAAGCTCAATTCATGATGAGCAATCAAAAACATGATCGGCA 600
Db 424 AATGAAAAATATCAAGTTCTGATTCGATTCGCAAAATTAATAATATCTTTCTGCA 483
OY 601 AAAGGATATGATAGTACAGCAACTTAATAGATTAGATGAGTAGTGCGCAGTG 660
Db 484 AAAGG-----CCTGAGCGTTTGGGACAGGCTGGCCATTA 516
OY 661 CAAAACGCTGATGATCGGCGCAATATATCATGATATCATGCTGCTCCGCTTACA 720
Db 517 CAAAACGCTGAGGAGCTGCGCAAACTATACGCGCTACACATGCTGCTTTGCAATAGCC 576
OY 721 GGTGACCCAAAAACAGTATGATCTGCACTTCAATTTATCTATCAAAAAGTGGTAT 780
Db 577 GGAGATCTCTAAAAATGCGGATGACACATCGATTTACATGTTCTATCAAAAAGTGGGCA 636
OY 781 ACATGATTTGACAGCTGGAATAATGCTGGAAGATATTGAAGATATGATTAATTTGT 840
Db 637 ACTTCTATTTGACAGCTGGAATAATGCTGGAAGATATTGAAGATATGATTAATTTGT 840
OY 841 CCAATGATCCGTATCTTAATATCAACACAGAGTGTGCTGCTGCTACTTAAC 900
Db 697 GCAATGATTTCTATCTTAATATCAACACAGAGTGTGCTGCTGCTACTTAAC 900
OY 901 AAAGTGGCCAAAGTGTGCTTATCTATCTATCTATCTATCTATCTATCTATCTATCT 960
Db 757 TGTGAGGAAAAATCCGTTTATCTATCTATCTATCTATCTATCTATCTATCTATCT 806
OY 961 ACCGCTGCTGTATCAAAATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1020
Db 807 -----CGGCAAAACACTGCAACTGCAACTGCAACTGCAACTGCAACTGCAACTG 858
OY 1021 GCTACACTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 859 TCT---TTGACATATACGCTGTAGAGATTTAATCAATCTTTGA---CGGTACGGA 912
OY 1081 ACAGTTTATCAAAATATTCAGCAATTTTCATGATGATGATGATGATGATGATGATGAT 1140
Db 913 AAAAGCTTACAAATATGTAAGCAATTCATGATGATGATGATGATGATGATGATGAT 972
OY 1141 CATACTTTAAGAGACCTCTCATATGTTGAGATTAAGGCGCATTAATATCTTTGGA 1200
Db 973 CATAGCGTGAAGATCTCTACATCTAGATTAAGGCGCATTAATATCTTTGGA 1032
OY 1201 GCGATTTACTGGAACAAGATGTTATGAGGCGATGCTGCTTCAATTAATTAACCTTAC 1260
Db 1033 GCAAACTCTGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1092
OY 1261 TATGCGGAAGTGAAGTCTTCTCCAGATGAATAAAATTAACCTCTTCAAGTCTTAA 1320
Db 1093 TATGCAAAAGCACAATCAATTTTCCGTCAGAAAGTCAAAAACCTTCGCAAAAGCATTAA 1152
OY 1321 AAACAAATGCTTCTTTAGCGAATGCTGATTTAGCATTTGTAATGGCGCATGACTAT 1380
Db 1153 AAACCAAGCGCTGATGATGAACAAGCGCTCTCGTATGATGATGATGATGATGATGAT 1212
OY 1381 ACAGTGAAGAAGTGTATGAACCATTTAGTGCATGAACAAGTGCATGATGATGATGAT 1440
Db 1213 ACATGAAAAAGATATGAACCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1272
OY 1441 CGCGCAATATATTTAAATGAATTAATGAATGATATCTATCTATCTATCTATCTATCT 1500
Db 1273 CGCGCAAGCGCTTTAAATGAAGCAAGTATGATGATGATGATGATGATGATGATGAT 1332
OY 1501 AAATGACAGTGTATGATTAACGACAAAGATGTTTATGATGATGATGATGATGATGAT 1560

```

```

Db 1333 AAAATGACGATTGACGCGATTAAGTATTAACATGCTGTTGTTTCTAAT 1392
OY 1561 TCCTTAATGGCCCAACCAACCCGATTAATTAACCTGACTGATGATGATGATGATGAT 1620
Db 1393 TCTTTAACTGGCCCAATCAACCGCTGAACAAACGCTGCTGTTTAAATGATCTT 1452
OY 1621 GACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1453 GATCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
OY 1681 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1513 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
OY 1741 CTGCGGCAAGCTTGGGCTTAATTAATTAAGGCTGACATCTGAGAGGAAATAG 1799
Db 1573 TTTGGCGCAAGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1631

```

RESULT 5

AA257330 standard; DNA: 1722 BP.

AA257330;

03-APR-2000 (first entry)

SacB and cpy chimeric gene #2.

Levansucrase; SacB; cpy; salt tolerance; vacuole guide peptide;

carboxypeptidase A; chimeric gene; transgenic plant; yeast;

Bacillus subtilis 168; Saccharomyces cerevisiae x8;

drought resistance; ds.

Chimeric - Saccharomyces cerevisiae.

Chimeric - Bacillus subtilis.

CN1231337-A.

13-OCT-1999.

08-APR-1998; 98CN-0101336.

08-APR-1998; 98CN-0101336.

(GENE-) INST GENETICS CHINESE ACAD SCI.

Chen S, Zhang H, Dong W;

WPI; 2000-087902/08.

Method for improving plant salt resistance using transgenic technology

Example: Fig 4; 12pp: Chinese.

A method has been developed for raising salt tolerance of plants using transgenic technology. The method comprises: cloning levansucrase gene SacB and vacuole guide peptide (cpy) sequence of carboxypeptidase A from *Bacillus* and DNA of yeast, using the two kinds of gene to create a chimeric gene, and using the chimeric gene to structure plant expression carrier, using the obtained plant expression carrier to transform a plant and screen for resistant seedlings. The method can obtain drought-resistant salt-tolerance plants, and the polymerase chain reaction (PCR) and Northern analysis of transgenic plants shows that the exogenous gene is integrated in the salt-tolerance transgenic plant. The present sequence represents a SacB and cpy chimeric gene from the present invention.

Sequence 1722 BP: 590 A; 390 C; 339 G; 403 T; 0 other;

Query Match 24.0%; Score 577.8; DB 21; Length 1722;

Best Local Similarity 66.4%; Pred. No. 5.3e-153;

Matches	916; Conservative	0; Mismatches	412; Indels	51; Gaps	4;
Qy	421 GTAGAGCGGACGACCCATATTTTTCGCGCAAAATGAAACAGCGGAGCTACAAAGAA				480
Db	335 GCAAGTGAACATATCAGCTTCGTCTCAATTAAGAAAGAAAGCAAAAGCCATATTAAGAA				394
Qy	481 GACTATGGTTTGGCCATATTTATACAGCGGCTGACATGCTTAAATAATTCAGAACAAAC				540
Db	395 ACATACGCGATTTCCATATTTACACGCGATGATATGCTGCAATCCCTGAAACAGCAAAA				454
Qy	541 AGTCTCAATTTAAAGTCCCTCATATTCATATCAGCAATCAAAAACATGATTGGCA				600
Db	455 AATGAAAAATATGCAATTCCTGAAATTCGATTCACAAATTAATAATATCTCTCTGCA				514
Qy	601 AAAGGTATGATTAAGTACGCACTTAATAGATTAGATGATGAGATAGCGGCACTG				660
Db	515 AAAAG-----CCTGGACGTTTGGGACACGCTGGCCATTA				547
Qy	661 CAAAACGCTGATGCTACTGCGCAAAATATATCATATCATATGCTCCGCTTTAGCA				720
Db	548 CAAAACGCTGACGCGACTGTCGCAAACTATACAGGCTACCATGCTCTTGCAATTAGCC				607
Qy	721 GGTACCCCAAAAACAGTATGATATCTCACTTATTTCTATCAAAAAGTCGCTGAT				780
Db	608 GGAATCTCTAAATATGGGATGACACATCGATTTACATGTTCTATCAAAAAGTCGCGCA				667
Qy	781 ACATCATATGACGCTGGAATAATGCTGAGAGATATTTGAAGATATGATATAATTTGTT				840
Db	668 ACTTCTATATGACGCTGGAATAATGCTGAGAGATATTTGAAGATATGATATAATTTGTT				727
Qy	841 CCAATATGATCCGTATCTTAATATCAACACAGAGATGTCAGTCTGCTACTTTAAC				900
Db	728 GCAATATGATCTATCTTAATATCAACACAGAGATGTCAGTCTGCTACTTTAAC				787
Qy	901 AAAGATGGCCAAATGCTTATCTATATGATATCTAGTATATCCGGAAGATGTTGA				960
Db	788 TCTGACGGAATAATGCTTATCTATATGATATCTAGTATATCCGGAAGATGTTGA				837
Qy	961 ACCGCTGCTGTAACCAATATCTTCACTGCTCAAGTAATATCCACGCGATGCA				1020
Db	838 -----CGCCAAACAAACGCTGACACATGCAACATGCAATGCAATGCAACG				889
Qy	1021 GCTACACTTAAAGTCATGATGATATCTGATATATATCTGCTTGTGATGCGGACGCT				1080
Db	890 TCT---TTGAACATCAACGCTGATGAGATATTAATCAATCTTTGA---CGGTACGGA				943
Qy	1081 ACAGTTATCAAAATATATCAATATGATATGATGATGATGATGATGATGATGATGATGAT				1140
Db	944 AAAACGTTACAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT				1003
Qy	1141 CATACCTTAAAGAGACCTCATATGATGATGATGATGATGATGATGATGATGATGATGAT				1200
Db	1004 CATACGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT				1063
Qy	1201 GCGAATCTGACACACATGATGATGATGATGATGATGATGATGATGATGATGATGAT				1260
Db	1064 GCAACACTGGAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT				1123
Qy	1261 TATGCGGAAGTGAAGCTCTTCCAGAAATGAAAAAATAAGCTTCAAGATCTTAA				1320
Db	1124 TATGCGCAAAAGCATATCTTCTCCGTCAGAAAGATCAAAATCTTCCAGAAAGCAT				1183
Qy	1321 AAACAAATGCTCTTAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT				1380
Db	1184 AAACGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT				1243
Qy	1381 ACATGAAAAAGTGTATGAAACATGATGATGATGATGATGATGATGATGATGATGAT				1440
Db	1244 ACATGAAAAAGTGTATGAAACATGATGATGATGATGATGATGATGATGATGATGAT				1303
Qy	1441 CGCGCAATATATTTAAATGAATATATGATGATGATGATGATGATGATGATGATGAT				1500
Db	1304 CGCGCAATATATTTAAATGAATATATGATGATGATGATGATGATGATGATGATGAT				1363

Qy	1501 AAAATGACGATGATGATTAACGCAAAAGATGTTATATGATGATGATGATGATGATGAT	1560
Db	1364 AAAATGACGATGATGATGATTAACGCAAAAGATGTTATATGATGATGATGATGATGAT	1423
Qy	1561 TCCTTAAATGGCCACACACACCCGATTAATGAAATGATGATGATGATGATGATGAT	1620
Db	1424 TCCTTAAATGGCCACACACACCCGATTAATGAAATGATGATGATGATGATGATGAT	1483
Qy	1621 GACCTGCTGATCTACACACACTTCTATGATGATGATGATGATGATGATGATGAT	1680
Db	1484 GATCCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1543
Qy	1681 AATGCTGATCTACACACACTTCTATGATGATGATGATGATGATGATGATGAT	1740
Db	1544 AATGCTGATCTACACACACTTCTATGATGATGATGATGATGATGATGATGAT	1603
Qy	1741 CTGCGGACAAAGCTTGGGTTAATTAATGAGGCTGACACATGAGAGAAATAG	1799
Db	1604 TTTGCGCAAGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGAT	1662

RESULT 6
 AAS68687 standard; cDNA: 2568 BP.
 AAS68687:
 13-FEB-2002 (first entry)
 DNA encoding novel human diagnostic protein #4491.
 Human: chromosome mapping: gene mapping: gene therapy: forensic:
 food supplement: medical imaging: diagnostic: genetic disorder: ss.
 OS Homo sapiens.
 WO200175067-A2.
 11-OCT-2001.
 30-MAR-2001: 2001WO-US08631.
 31-MAR-2000: 2000US-0540217.
 23-AUG-2000: 2000US-0649167.
 (HISE-) HISEQ INC.
 Drmanac RT, Liu C, Tang YT;
 WPI: 2001-639362/73.
 P-PSDB: ABG04500.
 New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity.
 Claim 1: SEQ ID No 4491: 103pp: English.
 The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC gene supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS6197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPD
CC at ftp.wipo.int/pub/published_pcl_sequences.

xx Sequence 2568 BP: 827 A; 582 C; 615 G; 544 T; 0 other:

Query Match 23.4%; Score 564.2; DB 23; Length 2568;
Best Local Similarity 64.7%; Pred. No. 4.7e-149;
Matches 962; Conservative 0; Mismatches 468; Indels 57; Gaps 6;

```

OY 357 GACGATGACCTTAAAGATGGCGAAAAAAGCAGCTCCCTTAACCTTCAGAGTGTAT 416
DB 3 GAACATGAACATCAAAAAAATTTTAAACAGCCAGCTTCTGACTTTTACGACTGCAC 62
OY 417 ATTACTAGAGCGGAGCGACCGCATATTTTGGCAGCAAAATGAACAGCGGAGCTACA 476
DB 63 TCTGGCAGGA---GGAGCGACTCAAGCCTTCGCCAAAGAAATTAACAAAAAGCATACA 119
OY 477 GGAACATATGTTTGGCCATATTACAGCGGCTACATGCTTAAATTTCCAGACACA 536
DB 120 AGAAACGTAAGCGGCTCTCATATTACAGCCATATATGCTGAGATCCCTAAACACGA 179
OY 537 AAACAGTCTCAATTTAAAGTGCCTCAATTCATGCATCAGCAACAAACATGATGTC 596
DB 180 GCAAAACGAAAAATACCAAGTGCCTCATTCATCAATCAACGATTAAATAATTTGAGTC 239
OY 597 GCGAAAGGGTATAGTACAGCACTTAATAGATTAGTATGAGTATGAGTATGAGTGC 656
DB 240 TGCAAAAG-----ACCTGATGTGTGGAGACAGCTGCC 272
OY 657 ACTGAAACGCTGATGTGCTACGCGCAATTAATCATGATATCAATCCTGCTCCGCTT 716
DB 273 GCTGAAACGCTGAGGAGCAAGTACGTAATACAGCGCTATACAGCTGTGTTGCTCT 332
OY 717 AGCAGTACCCAAAAAAGCAGTATGATATCTCACTTATTTATTCATCAAAAAGTGG 776
DB 333 TGGCGGAGCCCGAAGAGCGTATGATACATCATCATGTTTATCAAAAAGTGG 392
OY 777 TGATACATGATTCAGCTGGAAGAAATCTGGAAGATTTGAATATGATTAAT 836
DB 393 CGACACATCAATCGACGCTGGAAGAAAGCGGCGCTGTTTAAAGACAGCATTAAT 452
OY 837 TGTTCGAATGATCCGTATCTTAATATCAACAGAGAGTGTGAGTTCGTACTT 896
DB 453 CGAGCGCAACGATCCGATCTGAAGATCAGACGCAAGAAATGTCGGTCTGCAACCT 512
OY 897 AACCAAGATGCCAAGTCCGTTTATCTATACAGATTACTACGATTAATCCGAGATGG 956
DB 513 TACATCTAGCGAAAAATCCGTTTATCTACACTACATTCGGTAAACATTA----- 566
OY 957 TGGAAACGGTGTGTAACCAATCATTTCAACTGCTCAAGTAACTTATCCACCGGA 1016
DB 567 -----CGGCAAAACAAAGCCTGACAAACAGCAGAGTAATGTCTAAATCTGA 614
OY 1017 TGCAGCTACCTTAAGTCGATGAGTATCTGATATTAATCTGCTTGTGATGGGAG 1076
DB 615 T-----GACACCTCAAAATCAAGGAGTGAAGATACAAAAGATTTTGA---CGGAGA 668
OY 1077 CGGTACACTTATCAAAATATTTACAGCAATTTATCGATTAAGCAAGTGTGATTCAGTGA 1136
DB 669 CGGAAAAACATATCGAAGCTTACAGCACTTATGATGAAGCAATTAATACCTCGGGA 728
OY 1137 TAACCATCTTTAGAGACCTCACTATGTTGAAGATAAGGGCCATAAATATCTTGTCT 1196
DB 729 CAACCATAGCGTGAAGACCTCACTACGTGTAAGACAAAGCCATTAATACCTGTAT 788
OY 1197 TGAACGGAATCTGAACACAGATGTTTCAAGGATCGTCTTCAATTAATTAAGC 1256
DB 1111 11 11 11111 1 11 11 1111111 11111 111 11111
```

```

DB 789 CGAAGCCACACAGGGAAACGAAAAAGATACCAAGCGGAGAGATCTTTATTTAAACAAAGC 848
OY 1257 TTACATGCGGAGAGTACGCTTCTTTCACATGATAAAAAATAAAGCTTCAAGTCC 1316
DB 849 GTACTACGCGCGCGGACGAACTTCTTCCGTAAAGAAAAAGCAGAAAGCTTCAGCAGACGC 908
OY 1317 TAAAAACAAATTCCTTCTTAGCGAATGTGATTTAGGCAATTTGTAATGGCCGATGA 1376
DB 909 TAAAAAACGCGATGCTGAGTATAGCAAGCGGCGCCCTGGTATCATAGTAAATTAATGA 968
OY 1377 CTATACAGTAAAGTGTATGAAGCCATGTAGTGCATCAACACAGATAGCAGATGAAGT 1436
DB 969 TTACCATTTGAAAAAGTAATGAAGCGCTGATCATTTCAACACGGAAGTAACTGATGAAT 1028
OY 1437 CGAAGCGCCCAATATATTTAAATGAATGAATATGATATCTATTCAGATTCAGAGAG 1496
DB 1029 CGAGCGCGCGAATGTTTCAAAATGAAGCGCAAAATGTGACTTGTACTGATTCACGCGG 1088
OY 1497 ATCCAAATATGACAGTATGAATTAACGACAAAGATTTATATGCTAGGCGCCGAGG 1556
DB 1089 TTCAAAATATGACAGTATGATGATTAACCAAGCATATTTACATGCTTGTATGTATC 1148
OY 1557 CGACTCCTTAATGCGCCACACACACCCGATTAATGAATGAACTGTGATTAAGACATGA 1616
DB 1149 AAACCTTTAAACGCGCCTTACAAAGCGCTAACAACAAACAGGCGCTGTGCTCAAAATGG 1208
OY 1617 TCTTGACCCCTGATCTACACACACTTACTCTATGCGGTATCCCGCAACCTGAAG 1676
DB 1209 TCTTATGATCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
OY 1677 TAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1736
DB 1269 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1328
OY 1737 TCACCTGCGGAGCAAGCTTGGGGTAAATTAATGAAGGCTGACACATCTGAGAGGAGAA 1796
DB 1329 AACATTTCCGCAAGCTTCTTAATGAACATCAAAAGGCAATTAACATCCGTTGCAAAA 1388
OY 1797 TAGTTC---GGACAAAGCAATTCCTATAGCAGTATATCCCATTA 1840
DB 1389 CAGCATCTGAGAGAGACAGCAGCTGACATCATGATGATGATGATGATGATGATGATG 1435

RESULT 7
AAS68494
ID AAS68494 standard; cDNA; 1584 BP.
XX
AC AAS68494;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4298.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB: ABG04307.
```

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 1: SEQ ID No 4298; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1584 BP; 539 A; 358 C; 350 G; 337 T; 0 other:

Query Match 23.4%; Score 563.4; DB 23; Length 1584;
Best Local Similarity 65.0%; Pred. No. 6.1e-149;

Matches 938; Conservative 0; Mismatches 451; Indels 54; Gaps 5;

QY 357 GAGCATGAACCTTCAAAAGATTGGCGAAAAAGACAGCTCCCTAACCTTCAGAGACTGTAT 416
DB 162 GACATGACATCAATCAAAAAATTTAAACAGCCACATCTTGACTTTAGACGACTGACT 221
QY 417 ATTAGTAGAGCGAGCGACCGCATATTTTGGCCAGCAAAATGAACACGCGGACTACAA 476
DB 222 TCTGGCAGGA---GGAGGACTCAAGCCCTTCGCGAAAAATAACCAAAAAGATACAA 278
QY 477 GGAAGACATATGTTTGGCCATATTACACGGCGTGACATGCTAAATAATTCAGACACAA 536
DB 279 AGAAGCTAGCGGCTCTCTCTATTACAGCCCATGATATGCTGCAGATCCCTAAACACCA 338
QY 537 AAGCAGTCCTCAATTTAAAGTGCCTCAATTCATGATCAGCAATCAAAAACATTTGATTC 596
DB 339 GCAAAAGCAAAATACCAAGTGCCTCAATTCATGATCAGCAATCAAAAATTTGATTC 398
QY 597 GGCAAAAGGATATAGTAGAGCACTTATAGATTTAGATGTATGGATAGCTGGCC 656
DB 399 TGCAAAAGG-----ACCTGATGTGTTGGGACAGCTGGCC 431
QY 657 ACTGCAAAAAGCGTATGCTAGCTGCGCAAAATTCATGATGATACATGCTCGGCTTT 716
DB 432 GCTGCAAAAAGCGTATGCTAGCTGCGCAAAATTCATGATGATACATGCTCGGCTTT 491
QY 717 AGCAGGTGACCAAAAAACAGTATGATTCACCTTCATTTATTCATCAAAAAGTCCG 776
DB 492 TGCGGGAAGCCGGAAGAGCGCTGATGACATCATCATGTTTTCAAAAGGTCCG 551
QY 777 TGATACATGCTATACAGCTGGAAAAATGCTGGAAGATATTGTAAGATATGATTAAT 836
DB 552 CGACAACTCATGACAGCTGGAAAAACGCGGCGCTCTTTTAAAGCACGATTAAGTT 611
QY 837 TGTTCAAATGATCGATCTTAAATATCAACAGAGAGGTCAGAGTTTGTGCTACTTT 896
DB 612 CGAGCGCAGCATCGATCTGAAAGATCAGACGCAAGATGTCGGTTCGTGCAACCTT 671
QY 897 AACCAAGATGGCCAAAGTCCGTTATTCATACAGATTACTCAGGTAACTCTGGAAGATG 956

DB 672 TACATCTACGGAATAATCCGTTTATCTTACACCTGACATATCCGGTAACATTA----- 725
QY 957 TGGAAACCGGTGCTGTGAACCAATATCTTCAATGCTCTCACTAACTTATCCAGCCGA 1016
DB 726 -----CGGCAAAACAAAGGCTTGACAAACAGCCGAGGTAAATGTCTCAAAATCTGA 773
QY 1017 TGCAGCTACCTTAAGTGCATGGAGTATGATCATTAATCTGCTTGTGATGGCGGGA 1076
DB 774 T---GACACGCTCAAAATCAACGAGTGAAGATCACAACAAAGATTTTGA---CGGAGA 827
QY 1077 CGGTACAGTTTAAATTAATTCAGCAATTTATGATGAAGGCAAGTGGATTTGAGTGA 1136
DB 828 CGGAAACCAATATGAGCAAGTTCACAGCTTTATGATGATGAAGCAATTTATACATCCGCGA 887
QY 1137 TACCATATCTTTAGAGACCTCTACTATGTTGAAGATAGGCGCATTAATATCTTGTCTT 1196
DB 888 CAACCATAGCTGAGAGACCTCTACTAGTTGAAGACAAAGCCATTAATATCTTGTATTT 947
QY 1197 TGAAGCGAATCTGGAACACAGATGTTATCAAGGCGATCAGTCTTCATTAATTAAGC 1256
DB 948 CGAAGCCACACGGAACAGAAAGGATACCAAGGCGAAGATCTTATTTAACAAAGC 1007
QY 1257 TTACTATGCGGAAGTGAAGTCTTCTTCCAGATGAAAAATAAATACGTTCAAAAGTCC 1316
DB 1008 GTACTAGCGCGGCGGACAGAACTTCTCCGTAAGAAAGCCAGAAAGCTTCAGCAGAGCCC 1067
QY 1317 TAAAAAACAAATGCTTTTATAGCAATGATGATAGCATTTGTTGATTTGATGCGCATGA 1376
DB 1068 TAAAAAACGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1127
QY 1377 CTATACAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
DB 1128 TTACATGATGAAAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1187
QY 1437 CGAAGCGCCCAATATATTTAAATGATTAATTAATGATGATGATGATGATGATGATGATGAT 1496
DB 1188 CGAGGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1247
QY 1497 ATCCAAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556
DB 1248 TTCAAAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1307
QY 1557 CGACTCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1616
DB 1308 AACTCTTTAAACGCGCTTTCACAGCGCTCAACAAACAGAGGCTTGTGCTGCAAAATGG 1367
QY 1617 TCTTGACCTCTGATCTACACACACTTACTCTCATTTGCGGTATCCCGACCCCTGAAGG 1676
DB 1368 TCTTGATCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1427
QY 1677 TAATATGTTGTTACTCACAAGTATATGACGAATAGAGGCTTCTATCCAGAAATCACTC 1736
DB 1428 CAACAATGTTGTTTCAACAACTACATGACAAACAGAGGCTTCTGAGAGTAATAAAGGC 1487
QY 1737 TCACCTGCGGGAACAAGTGTGGGTTAATTAATAAGGCTGACACATCTGAGAGGAGAAA 1796
DB 1488 AACATTTGCGCAAGCTTCTTAATGATCATCAAGGCAATTAACATTCGTTGCAAAA 1547
QY 1797 TAG 1799
DB 1548 CAG 1550
RESULT 8
AAS85834
ID AAS85834 standard; cDNA; 2044 BP.
XX AAS85834;
XX AC
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #21638.
DE

Db 1888 CAACAATGTGGTTATCACAAGCTACATGACAACAAGAGCGTCTTCGAG3ATMAAAGGC 1947
QY 1737 TCACCTGGGGGACACAGCTTGGGGTTAATTTAAAGGCTCTGACACATCTGAGGAGAAA 1796
Db 1948 AACATTGGCGGACAGCTTCTTAATGAACATCAAAAGCAATAAACATCC3TTGTCAAAA 2007
QY 1797 TAG 1799
Db 2008 CAG 2010

RESULT 9
AAS87384/C
ID AAS87384 standard; cDNA: 2044 BP.
AC AAS87384;
XX 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #23188.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR P-PSDB; ABG23197.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS
PS Claim 1; SEQ ID No 23188; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (I) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcf_sequences.

Sequence 2044 BP; 426 A; 438 C; 506 G; 674 T; 0 other;

Query Match 23.4%; Score 563.4; DB 23; Length 2044;
Best Local Similarity 65.0%; Pred. No. 7e-149;
Matches 938; Conservative 0; Mismatches 451; Indels 54; Gaps 5;

QY 357 GACGATGACCTTCAAAAGATTGGCGAAAAAGACAGTCCCTTAACCTTCAGAGCTGTAT 416
Db 1423 GAACATGAACATCAAAAAAATTGTAAACAGCCACATCTTGACTTTAGGACTGCACCT 1364
QY 417 ATTAGTAGAGCGGACGACCGCATATTTTGGCAGCAATGAACAGCGGGACTACAA 476
Db 1363 TCTGGCAGCA--GGAGGACTCAAGCCCTTCGCAAAAGAAATAAACCAAAAGCATACAA 1307
QY 477 GGAAGCATGTGTTTGGCCATATTAACAGCGGCTGACATGCTAAATAATCCAGGACCA 536
Db 1306 AGAAAGCTAGCGCGCTCTCATATTACAGCGCATGATATGCGAGATCCCTTAACACCA 1247
QY 537 AAACAGTCCGTAATTTAAAGTGCCCTCAATTCATGCAATCAGCAATCAAAACATTGATTC 596
Db 1246 GCAAAAGCAAAATATCAAGAGGCTCATTTGATCAATCAACGATTAATAATATTGAGATC 1187
QY 597 GCGAAAGGGTATGTAAGTCAGGCACTTAATGATTTAGATGTATGGATPAGCTGGCC 656
Db 1186 TGCAAAAGG-----ACCTGATGTGGGACAGCTCGCC 1154
QY 657 ACTGCAAAAGCGTGATGTAGTCGGGCAATTTATCATGGATATCATCGTCCGCTTT 716
Db 1153 GCTGCAAAAGCGTGAGGAGACAGTGTGAATACAGGCTATACGTTGTGTGCTCT 1094
QY 717 AGCAGGTAGCCCAAAAAACAGTATGATACCTCCTTATTTATCTATCAAAAGTCCG 776
Db 1093 TCGGGGAAGCCCGAAGAGCGGTGATGACACATCATCATGTTTATCAAAAGCTCG 1034
QY 777 TGATCATCGATTGACAGCTGGAAAAATGCTGGAAAGTATTGAAATATGATTAATT 836
Db 1033 CGACAACTCAATCGACAGCTGAAAAACGCGGCGCTCTTTAAAGACAGGATTAAGTT 974
QY 837 TGTCCAAATGATCGGTATCTTAATATCAAAACAGAGAGTGGAGTTCGCTACTTT 896
Db 973 CGAGCCAAACGATTCGATCTGAATGATCAGACGAAAGATGTCGGTCTGCAACCTT 914
QY 897 AACCAAGATGGCCAGTCCGTTTATCTATACAGATTACTCAGTAATCCGTAAGATGG 956
Db 913 TACATCTGACGGAATAATCCGTTTATCTAGCTACATTTCCGTAACATTA----- 860
QY 957 TGGAAACGGTGTGTTAACCATAATCATTTCACTGCTCAAGTAACTTAATCCACCGCA 1016
Db 859 -----CGGCAAAACAAAGCCTGACACAGCGAGTAAATGTGTCAAAATCTGA 812
QY 1017 TGCAGTACACTTAAGTCGATGAGTATCTGATCATTAATCTGCTTGTGATGGCGGAGA 1076
Db 811 T---GACACACTCAAAATCAACGGAGTGAGATACAAACGATTTTGA---CGGAGA 758
QY 1077 CGGTACAGTTTATCAAAATATTACAGCAATTTATCGATGAAGCAAGTGGATTTCAGTGA 1136
Db 757 CGGAATAACATATCGAAGCTTACAGAGTTATGATGAAGCAATTTATACATCCGGCA 698
QY 1137 TAACCATCTTAAAGACACCTCACTATGTGTAAGATTAAGGCCATAATATCTTGTCT 1196
Db 697 CAACCATATCGGTGAGAGACCCCTCACTAGCTGAAGAACAAAGCCATAATATCTGTAT 638
QY 1197 TGAACGAATCTGGAACAACAGATGTTATCAAGCCATGATCTTTCATTAATAAAGC 1256
Db 637 GCAAGCCAAACAGGGAAGAAAAGGATACCAAGGCAAGATCTTTATTTAATAAAGC 578
QY 1257 TTACTATGCGGAATGAGTCCGTTCTTCCAGATGAATAAATAAATACGTTCAAGTCC 1316
Db 577 GTACTAGCGCGGCGGACGAGACTTCTTCCGTAAGAAAAGCAGAAAGCTTACGAGAGCC 518
QY 1317 TAAAAACAAATGTCTTTAGCAATGTGATTTAGCATTTGTTGATTTGGCCGATGA 1376
Db 517 TAAAAAAGCGATGCTGAGATTAGCGAAGCGGCGCCCTCGTATCATAGATTAAATAATGA 458


```

QY 1077 CGGTACAGTTATGCAAAATATTCAGCAATTATGATGAAGCAAGTGGATTTCAGGTGA 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 669 CGGAAAAACATATGAGAAAGCTTCAGCAAGTTATGATGAGCAAGCAATTATCATCCGGGGA 728
QY 1137 TAACCATATCTTAAAGACACCTCAGTATGTTGAAGATPAGGCGCAATAAATCTTGTCTT 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 729 CACCAATACGCTGAGAGACCTCAGTATGTTGAAGCAAGCAAGCAATAAATCTTGTATAT 788
QY 1197 TGAAGCGAATATCGAACAACAGATGTTATGACGAGGATGAGCTTTCATATTAAGC 1256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 789 CGAAGCCAGACGAGGACAGAAAGCGATACCAAGAGCAAGCAATCTTATTTAAACAAAGC 848
QY 1257 TTACTATGCGGGAAGTGAAGCTCTTCTTCAGAAATGAAAAAAATTAAGCTTCTTAAGTCC 1316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 849 GTACTAGGCGGGGCGACAGACTCTCTCCGTAAGAAAGCAAGAGCTTCAGCAGAGCGC 908
QY 1317 TAAAAAACAAATTCCTCTTATGCGAATGTGCAATPAGCAATCTTGAATGGCCGATGA 1376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 909 TAAAAAACGCGATCTGATGAGCAAGCGCGCCCTCGATATCAATAGATTAATTAATGA 968
QY 1377 CTATACAGTGAAGTGTATGAAACGATTAAGTGCATCAACACAGTAGAGAGATGAAGT 1436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 969 TTACACATTTGAAAAAGTAAATGAAGCCGCTGATCATCTTAACACAGGTAAGTGAAT 1028
QY 1437 CGAAGCGCGCAATATATTTAAATGAATTAATTAATGATCTATTCACGAGTTCAGAGG 1496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1029 CGAGCGCGCGAATGTTTCAAAATGAAGCAAGCAATGACTGTTGCACTGATTCACGCGG 1088
QY 1497 ATCCAAATGACGAGTGTGAATTAAGCAACAAAGTGTATATGCTAGGCGCGGAGG 1556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1089 TTCAAAATGACGATCGATGTTAATTAAGTCAACAGATTTAATGCTTGTATGATTC 1148
QY 1557 CGATCTCTTAATGAGCCACACACCGCATTAATGAAGTGCATGATTAATGAACATCA 1616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1149 AATCTCTTAACCGGCCCTTAACAGCCGCTGATCAACAAACAGGCTTGTGCTCAAAATGG 1208
QY 1617 TCTTACCCCTGCTGATCTCAGACACAGTACTCTATCTATGCGGTATCCGCAAGTGAAG 1676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1209 TCTTGATCAAAAGATGATGATCATCTACTCTCACTTCCAGTGGCGCAAGCAAAAGG 1268
QY 1677 TAATTAATGTGTACTCAGCAAGTATATGACGAATAGAGGCTTCTATCCAGATCACTC 1736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1269 CAAACAAATGTGTATTCACAAAGCTACATGACAAACAGAGGCTTCTCGAGATTAAGGCG 1328
QY 1737 TCACCTGCGGAGCAAGCTTGGGTTAATTAAGGCTGACACATCTGAGAGAGAAA 1796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1329 AACATTTGCGCCAGCTTCTTAATGAACATCAAAAGGCAATTAACATCCGTTGCAAAA 1388
QY 1797 TAG 1799
    ||
Db 1389 CAG 1391

```

```

RESULT 11
AAS75941
ID AAS75941 standard; cDNA; 2325 BP.
AC AAS75941;
XX
XX 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #11745.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX W0200175067-A2.
XX
XX 11-OCT-2001.
XX

```

```

PF 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HXSE-) HXSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI WPI: 2001-639362/73.
DR P-PSDB: ABG11754.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11745; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
SQ

```

```

SQ Sequence 2325 BP; 713 A; 593 C; 566 G; 453 T; 0 other;
Query Match 23.4%; Score 563.4; DB 23; Length 2325;
Best Local Similarity 65.0%; Pred. No. 7.5e-149;
Matches 938; Conservative 0; Mismatches 451; Indels 54; Gaps 5;
QY 357 GAGCATGACTTCAAAAGATTGGGCAAAAAGACAGCTGCCATACTTCAGAGCTGCTAT 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 GAACATGAAACATCAAAAATTTGTAAMACAGCCAGTCTGACTTTAGACTGCAC 62
QY 417 ATTAGTAGAGACGCGACGCGCATATTTTGGCCAGCAATGAACAGCGGGAGCTACAA 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 TCTGCGCAGCA--GGAGCGACTCAAGCCTTGCGCAAAAATAACCAAAAAGCATACAA 119
QY 477 GGAAGACTATGTTTGGCCATATTACAGCGGCTGACATGCTTAATAATTCAGAGCAACA 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 AGAAGCTAGCGCGCTCTCTCATATTACAGCGCATGATATGCTGAGATCCCTTAACAGCA 179
QY 537 AAACAGTCCATTAATTAAGTGCCTCAATTCATGATGACCAATCAAAAATTTGATTC 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 GCAAAAGCAAAAATACCAAGTGCCTCAATTCGATCAATCAAGATTAATAATATTGAGTC 239
QY 597 GCGAAAAGGATATGATAGTCAAGCACTTAATGATTTAGATGATGAGGATAGCTGCC 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 TGCAAAAG-----ACTGATATGTGGGACACCTGGC 272
QY 657 ACTGCAAAAGCTGATGATGCGGCAATTAATCATGATGATCATATCGCTCGGCTT 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 GCTGCAAAAGCTGATGAGCAAGTACGATGATACAGCGCTATACGTTGTTGCTCT 332
QY 717 AGCAGGTGACCCAAAAAACAGTATGATCTCAGCTTCAATTAATTCATCAAAAAGTCGG 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      860  GAACATGAACTCAAAAAATTGTAAGAACCAAGCTCTGACTTTTACGACTGCACCT  919
Oy      417  ATTAGTAGAGCGGAGCGACCGCATATTTTGGCAGCAAAATGAGACAGCGGGAGTACAA  476
Db      920  TCTGGCAGAGA---GAGGAGACTCAAGCCTTGGCGAAAGAAAATATACCAAAAAAGCTACAA  976
Oy      477  GGAAGACTAGTGGTTTGGCCATATTACAGCGCGTGACATGCTTAAAAATTCAGAGCAACA  536
Db      977  AGAAACGTACGGCGCTCTCATATTACAGCCCATGATATGCTGAGATCCCTTAACACAGA  1036
Oy      537  AAACAGTCCCTCAATTAAAGTGGCCCATTTGATGATCAGCATCAGCAATCAAAAACATTGATC  596
Db      1037  GCAAAACGAAAAATACCAATGAGTCCCTCAATTCATCAATCAAGATTTAAATATTTAGATC  1096
Oy      597  GCGAAAAAGGTATGATAGTCAGCGCACTTAATGATTAGATGATGATGATGATGATGATG  656
Db      1097  TGCAAAAGG-----ACCTGATGTGTGTGGTACAGCTGCGCC  1129
Oy      657  ACTGCAAAACGCTGATGCTGCTGCTGCGGCAAAATATCATGATATCATCATCTCTCCGCTTT  716
Db      1130  GCTGCAAAACGCTGACGAGACAGTGTGAAATACAAAGGCTATCAGCTTGTGTTGCTCT  1189
Oy      717  AGCAGGTGACCCCAAAAAACATGATGATCTCCACTTCAATTTATTTCTATCAAAAAGTGG  776
Db      1190  TCGGGGAAGCCCGAAAGCGCTGATGACATCAATCTACATGTTTTTATCAAAAAGCTGG  1249
Oy      777  TGATACATCGATTGACAGCTGGAATAATCTGGAAGAGATTTTGAAGATATGATTAAT  836
Db      1250  CGACACATCTCATTCGACAGCTGGAATAATCGCGGCGCTCTTTTAAAGACAGCGATTAAT  1309
Oy      837  TGTTCCAATGATCGATCTGTTAAATATCAAAACAGAGTGGTCTGCTACTCTTT  896
Db      1310  CGAGCCCAACGATCCGATCTGAAAGATCAGACGCAAGATATGTCGGTCTGCAACCTT  1369
Oy      897  AACCAAGATGGCCCAAGTCCGTTTATTTCTATACAGATTACTACAGTATCTCGAAGATGG  956
Db      1370  TACATCTGACGAAAAATCCGTTTATTTACACGACTGACTGCTTCCGTTAAACATTA-----  1423
Oy      957  TGAACCCGCTGCTGTAACCAAAATCATTTCAACTGCTCAAGTAACTTATCCACCGCGA  1016
Db      1424  -----CGGCAACAAAGCCTGACAAAGCGGAGGATATGCTCAAAATCTGA  1471
Oy      1017  TGCAGCTACACTTAAAGTCGATGAGATATCTGATCATAAATCTGCTTTGATGGCGAGA  1076
Db      1472  T---GACACACTCAAAATCAACGAGTGAAGATACAAAACGATTTTGA---CGGAGA  1525
Oy      1077  CGGTACAGTTTATCAAAATATTCAGCAATTTATGATGAGAGCAAGTGGATTTTCAGTGA  1136
Db      1526  CGGAAAAACATATCAAGAGCTTCAGCAGTTTATGATGAGGCAATTTATPCATCCGCGGA  1585
Oy      1137  TAACCATACTTTAAGAGACCTCACTATGTTGAAGATTAAGGCGCATTAATATCTGTCTT  1196
Db      1586  CAACCATGCGCTGAGAGACCTCACTACTAGTTGAAGACAAAGGCCATTAATACCTGTAT  1645
Oy      1197  TGAAGCGAATCTGACACACAGATGTTATCAAGCGCATGCTTTCAATATATTAAG  1256
Db      1646  CGAAGCCCAACGCGGAAACAGAAAACGATACCAAGCGCAAGATCTTTATTTAAACAAAGC  1705
Oy      1257  TTACTATGGGGAGTGAAGTACTCTTCCAGATCAAAAATAAATAACGCTTCAAGTCC  1316
Db      1706  GTACTAGCGCGCGGACGAAACTCTCTGTTAAAGCCAGAAAGCCATGACGACAGCGC  1765
Oy      1317  TAAAAAACAAATGCTCTTATAGCAATGATGATTAAGCATTTGTTGAATGGCCGATGA  1376
Db      1766  TAAAAAACGCGATGCTGAGTTAGCAAGCGCGCCCTCGTATCATAGAGTTTAAATATGA  1825
Oy      1377  CTATACAGTGAAGAGTGTATGAAGCAATTAAGTGCATCAAAACACAGTACAGATGAAGT  1436
Db      1826  TTACACATTTGAAGAAATGAAGCCCGCTGATCACTCTCAAAACAGGTAACTGATGAAT  1885
Oy      1437  CGAAGCGGCAATATTTTAAATGAATAATTAATGATATCTATTCAGAGATTTCAAGAG  1496
Db      1886  CGAGCGCGCAATGTTTCAAAATGAACGCAAAATGTTACTGTTCACTGATTCAGCGCG  1945

```

```

Oy      1497  ATCCAAATGACAGTGAATGAATTAACAGCAAAAGTGTATATAGCTAGGGCCGAGG  1556
Db      1946  TTCAAAATGAGAGATGATGATTAATTAACATATTTACATGCTTGGTATGATTC  2005
Oy      1557  GCATCTCTTAATATGCGCCACACACACCCGATTAATGAATGAATGATGATGATGAT  1616
Db      2006  AAACCTTTTAAACCGCCCTTACAAAGCCCTGAAACAAACAGGCGTGTGCTCAATGG  2065
Oy      1617  TCTTGACCTGCTGATCTCACACACACTTACTCTCATTTGGCGTATCCCGACCTGAGG  1676
Db      2066  TCTTGATCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG  2125
Oy      1677  TATATATGCTGATCAACAGTATATGACGAATGAGCTCTTATCCAGAACATCACTC  1736
Db      2126  CAACATGATGATTAACAGAGTACATGACATGACAAAGAGGCTTCTTGAGGATTAAGAGC  2185
Oy      1737  TCACCTGCGGCAAGCTTGGGTTAAATTAAGAGTCTGACACATCTGAGAGAGAAA  1796
Db      2186  AACATTTGGCCCAAGCTTCTTAAATGAACATCAAGCAATTAACATCCGTTGCAAAA  2245
Oy      1797  TAG 1799
Db      2246  CAG 2248

RESULT 13
AAS69133
ID AAS69133 standard. cDNA: 2370 BP.
XX
XX AAS69133;
XX
XX 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4937.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX P-PSDB: ABG04946.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations.
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity
XX
XX Claim 1: SEQ ID NO 4937; 103pp: English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or as
XX quantitating a polypeptide in tissue, as molecular weight markers and as

```



```
Db      2158  TCTTGATCCAAACGATGTGACATTCACTCTGCGAGTGCGCGCAAGCAAAGG 2217
OY      1677  TATATGTGTACTCACAAGTTATATAGGAATAGAGGCTTCTATCCAGACATCCTC 1736
Db      2218  CAGCATTGTGTTATCACAAGCTACATGACAAACAGAGGCTTCTCGAGGATAAAAAGG 2277
OY      1737  TCACCTGCGGACCAAGCTTGGGGTTAATATTAAGGGTCTGACACATCTGAGAGAGAAA 1796
Db      2278  AACATTGCGCCAGCTTCTTATGAACATCAAGGCAATAAACATCCGTGTCAAAA 2337
OY      1797  TAG 1799
Db      2338  CAG 2340
```

Search completed: June 18, 2003, 03:10:34
Job time : 537.636 secs

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Run on: June 18, 2003, 01:50:51 ; Search time 6277.52 Seconds

11163.571 Million cell updates/sec

Title: US-09-986-682B-5

Sequence: 1 CGGGGAATACTAGATTCC.....TACCAATAAAATGCTGATC 2408

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

1:	gb_ba:*
2:	gb_htg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vi:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htg_mus:*
34:	em_htg_pln:*
35:	em_htg_rod:*
36:	em_htg_nam:*
37:	em_htg_vrt:*
38:	em_sy:*
39:	em_htg_om:*
40:	em_htg_mus:*
41:	em_htg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	2408	100.0	2408	1	AB010272	AB010272 Bacillus
2	2408	100.0	2408	6	AR078182	AR078182 Sequence
3	2408	100.0	2408	6	AR208488	AR208488 Sequence
4	2408	100.0	2408	6	E15383	E15383 Bacillus sp
5	1365	56.7	1365	6	AR078181	AR078181 Sequence
6	1365	56.7	1365	6	AR208487	AR208487 Sequence
7	1365	56.7	1365	6	E15382	E15382 Bacillus sp
8	756.6	31.4	11318	1	AE007686	AE007686 Clostridi
9	627	26.0	1449	6	AX434229	AX434229 Sequence
10	611.8	25.4	2007	1	BSSACBG	X02730 Bacillus su
11	611.8	25.4	5941	12	SYNPELB	X05083 Cloning vec
12	611.8	25.4	6347	12	SYNPELA	L05081 Cloning vec
13	611.8	25.4	6798	12	SYNPELB	L05082 Cloning vec
14	611.8	25.4	5297	12	AF048702	AF048702 Site-spec
15	611.8	25.4	88063	1	BS294043	Z994043 B.subtilis
16	611.8	25.4	209510	1	BSUB0018	Z994043 Bacillus su
17	610.4	25.3	5831	12	AF047518	AF047518 Cloning v
18	610.4	25.3	5842	12	AF004910	AF004910 Cloning v
19	610.4	25.3	5846	12	XXU17500	U17500 Cloning vec
20	610.4	25.3	6349	12	AF047519	AF047519 Cloning v
21	607.2	25.2	10317	6	AR162197	AR162197 Sequence
22	594.6	24.7	1422	1	BSU34874	X52988 Bacillus st
23	593.4	23.4	2350	1	BASACB	U80928 Cloning vec
24	563.4	23.4	11612	12	CVU08929	U09128 pacB/IT PI
25	563.4	23.4	16011	12	XXU09128	U09128 pacB/IT PI
26	563.4	23.4	16297	12	CVU75991	U75991 Cloning vec
27	563.4	23.4	18772	12	AF133437	AF133437 Cloning v
28	563.4	23.4	19500	12	CVU75992	U75992 Cloning vec
29	563.4	23.4	22530	12	AB020028	AB020028 Transform
30	563.4	23.4	107872	2	AC013485	AC013485 Trypanoso
31	563.4	23.4	110000	2	AC058788	Continuation (3 of
32	563.4	23.4	113316	2	AC007866	AC007866 Trypanoso
33	563.4	23.4	122526	2	AC012647	AC012647 Trypanoso
34	563.4	23.4	133928	2	AC091089	AC091089 Oryza sat
35	563.4	23.4	133901	2	AC008146	AC008146 Trypanoso
36	563.4	23.4	146349	2	AC007863	AC007863 Trypanoso
37	563.4	23.4	151443	2	AC007865	AC007865 Trypanoso
38	563.4	23.4	154413	2	AC007862	AC007862 Trypanoso
39	563.4	23.4	155204	2	AC007926	AC007926 Trypanoso
40	563.4	23.4	156128	2	AC113365	AC113365 Mus muscu
41	563.4	23.4	156709	2	AC013484	AC013484 Trypanoso
42	563.4	23.4	156754	2	AC008031	AC008031 Trypanoso
43	563.4	23.4	157774	2	AC013353	AC013353 Trypanoso
44	563.4	23.4	159717	2	AC008368	AC008368 Trypanoso
45	563.4	23.4	162249	2	AC009259	AC009259 Trypanoso

ALIGNMENTS

RESULT 1	AB010272	2408 bp.	DNA	linear	BCT 07-AUG-1996
LOCUS	AB010272				
DEFINITION	Bacillus sp. gene for Beta-fructofuranosidase, complete cds.				
ACCESSION	AB010272				
VERSION	AB010272.1	GI:3399690			
KEYWORDS	Beta-fructofuranosidase.				
SOURCE	Bacillus sp. (strain:V230) DNA.				
ORGANISM	Bacillus sp.				
REFERENCE	1 (sites)				
AUTHORS	Kurimoto, M.				
TITLE	Cloning and sequencing of beta-fructofuranosidase gene from				
JOURNAL	Bacillus sp. V230				
	Unpublished				

REFERENCE 2 (bases 1 to 2408)

Tsutsuki, K.
Direct Submission
Submitted (08-JAN-1998) Keiji Tsutsuki, Hayashibara Biochemical Laboratories, Inc., Amase Institute; 7-7 Amase-mihami machi, Okayama, Okayama 700, Japan (E-mail: amasehepo.harenet.or.jp, Tel: 086-231-6731, Fax: 086-231-6738)

FEATURES

source

1. 2408
/organism="Bacillus sp."
/strain="V230"
/db_xref="taxon:1409"
/note="Isolated from soil"
361. 1824
/codon_start=1
/transl_table=11
/product="beta-fructofuranosidase"
/protein_id="BAA32083.1"

CDS

BASE COUNT 777 a 492 c 496 g 643 t
ORIGIN
Query Match 100.0%; Score 2408; DB 1; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGGGGAAATCTAATTCATTCGCAATTCGCCAGCTTGGTGAAGAGAGCTTCG 60
1 CGGGGAAATCTAATTCATTCGCAATTCGCCAGCTTGGTGAAGAGAGCTTCG 60
61 GACTCCCATGCGAGTGTGCGTAAGGCAAGCTTACATGAGCAATTCATAGAGAT 120
61 GACTCCCATGCGAGTGTGCGTAAGGCAAGCTTACATGAGCAATTCATAGAGAT 120
61 GACTCCCATGCGAGTGTGCGTAAGGCAAGCTTACATGAGCAATTCATAGAGAT 120
121 TTCAAAAGTCTGTGGATTAAGGCAAGCTTGAATTAACCTTTGACGCTGTGTG 180
121 TTCAAAAGTCTGTGGATTAAGGCAAGCTTGAATTAACCTTTGACGCTGTGTG 180
121 TTCAAAAGTCTGTGGATTAAGGCAAGCTTGAATTAACCTTTGACGCTGTGTG 180
181 GGGGCTTTTGTGTTTATTTATTCACATGCAATGCTCCATCCCTATATCATTTTA 240
181 GGGGCTTTTGTGTTTATTTATTCACATGCAATGCTCCATCCCTATATCATTTTA 240
241 GACGAAATTTCAATCAATCCATGCGATCCCAATTAACCTGCTCTATACCTTTTA 300
241 GACGAAATTTCAATCAATCCATGCGATCCCAATTAACCTGCTCTATACCTTTTA 300
241 GACGAAATTTCAATCAATCCATGCGATCCCAATTAACCTGCTCTATACCTTTTA 300
301 TTAAATTAAGAACTATCAAGAGCTTCTTATCAATTCATATATTCAGAGGAGAG 360
301 TTAAATTAAGAACTATCAAGAGCTTCTTATCAATTCATATATTCAGAGGAGAG 360
361 ATGAACCTCAAAAGTTGGCGGAAAGAGAGCTGCGGTAACCTTGAAGAGCTGTATTA 420
361 ATGAACCTCAAAAGTTGGCGGAAAGAGAGCTGCGGTAACCTTGAAGAGCTGTATTA 420
421 GTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
421 GTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
481 GACTATGTTTGGCCATATTAACAGCGCTGACATGCTAATAAATTCAGAGCAACAAAC 540
481 GACTATGTTTGGCCATATTAACAGCGCTGACATGCTAATAAATTCAGAGCAACAAAC 540
541 AGTCTCAATTTAAGTGCCTCAATTCATGATGATGATGATGATGATGATGATGATG 600
541 AGTCTCAATTTAAGTGCCTCAATTCATGATGATGATGATGATGATGATGATGATG 600

541 AGTCTCAATTTAAGTGCCTCAATTCATGATGATGATGATGATGATGATGATGATG 600
601 AAAGGTTATGATTAATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
601 AAAGGTTATGATTAATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
661 CAAAGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
661 CAAAGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 GGTGACCCCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
721 GGTGACCCCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
781 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
781 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
841 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
841 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
901 AAAGTGGCCAAAGTCCGTTATTTATTCATGATGATGATGATGATGATGATGATGAT 960
901 AAAGTGGCCAAAGTCCGTTATTTATTCATGATGATGATGATGATGATGATGATGAT 960
961 ACCGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961 ACCGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021 GCTACACTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021 GCTACACTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1081 ACAGTTTATCAAAATATTCAGCAATTCATGATGATGATGATGATGATGATGATGAT 1140
1081 ACAGTTTATCAAAATATTCAGCAATTCATGATGATGATGATGATGATGATGATGAT 1140
1141 CATACCTTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1141 CATACCTTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1201 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1201 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1261 TATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
1261 TATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
1321 AAACAAATTTGCTTTTATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1380
1321 AAACAAATTTGCTTTTATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1380
1381 ACAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
1381 ACAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
1441 GCGGCAATTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
1441 GCGGCAATTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
1501 AAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
1501 AAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
1561 TCCTTAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
1561 TCCTTAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
1621 GACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1621 GACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

1681 AATGTGACTACACAGTTATATGACGATAGAGCTTCTATCCAGAACATCACTCTCAC 1740
1681 AATGTGACTACACAGTTATATGACGATAGAGCTTCTATCCAGAACATCACTCTCAC 1740
1741 CTGGGGGCAAGCTTGGGGTTAATTTAAAGGCTGACACATCTGAGAGGAAATAGT 1800
1741 CTGGGGGCAAGCTTGGGGTTAATTTAAAGGCTGACACATCTGAGAGGAAATAGT 1800
1801 TCCGACACAGGACATTCCTCAGGATATCTCCCATTAATAAGAAATGACAGGCA 1860
1801 TCCGACACAGGACATTCCTCAGGATATCTCCCATTAATAAGAAATGACAGGCA 1860
1861 GATCCTACCTTTTCCCAACCTTTTAAATAAGAGTTTGTACCACTTTAGAA 1920
1861 GATCCTACCTTTTCCCAACCTTTTAAATAAGAGTTTGTACCACTTTAGAA 1920
1921 GAAAAAAGATCTTGGCGCGGCGCAATTACCCATGCGAAGGAGCTACCCAAATAAAA 1980
1921 GAAAAAAGATCTTGGCGCGGCGCAATTACCCATGCGAAGGAGCTACCCAAATAAAA 1980
1981 AGAATTTTCCCTTTTGGGGGGGAGTCTTTTCTATCTGGGGTGGGGATTTGTG 2040
1981 AGAATTTTCCCTTTTGGGGGGGAGTCTTTTCTATCTGGGGTGGGGATTTGTG 2040
2041 GCGCCACACAGGAGCTTTATTTATTTATAGCATGATCTGGCGATTTTGCCTGGCTA 2100
2041 GCGCCACACAGGAGCTTTATTTATTTATAGCATGATCTGGCGATTTTGCCTGGCTA 2100
2101 CTATTATTCAGAGCGCGCCAGCAAAATATTAAACAGGAAATCTTAAGATAGCGCGT 2160
2101 CTATTATTCAGAGCGCGCCAGCAAAATATTAAACAGGAAATCTTAAGATAGCGCGT 2160
2161 ACAGTTTCCAAACACCTCTGAGGAGGATGATTCGCTCAATACACATTTCTGTC 2220
2161 ACAGTTTCCAAACACCTCTGAGGAGGATGATTCGCTCAATACACATTTCTGTC 2220
2221 CGGACAAATGAAAAATGATCCGAGAGGCCATCTTTTGGGGAGATATCATTA 2280
2221 CGGACAAATGAAAAATGATCCGAGAGGCCATCTTTTGGGGAGATATCATTA 2280
2281 TTACCTTACAAACAAGCTATCCAGATGAAATGATGCGAATGGCGGATGCACTC 2340
2281 TTACCTTACAAACAAGCTATCCAGATGAAATGATGCGAATGGCGGATGCACTC 2340
2341 CGAAGATTATTCGCTTGGAGGAGGAGGATTCGCTCCGAAATGATACCAATAAAA 2400
2341 CGAAGATTATTCGCTTGGAGGAGGAGGATTCGCTCCGAAATGATACCAATAAAA 2400
2401 TGGTGATC 2408
2401 TGGTGATC 2408

RESULT 2
AR078182 2408 bp DNA Linear PAT 31-AUG-2000
LOCUS AR078182
DEFINITION Sequence 5 from patent US 5962297.
ACCESSION AR078182
VERSION AR078182.1 GI:10004928
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2408)
AUTHORS Tsusaki,K., Kubota,M. and Chaen,H.
TITLE Polypeptides having .beta.-fructofuranosidase activity
JOURNAL Patent: US 5962297-A 5 05-Oct-1999;
FEATURES
1..2408
Location/Qualifiers
source /organism="unknown"
BASE COUNT 777 a 492 c 496 g 643 t
ORIGIN

Query Match 100.0%; Score 2408; DB 6; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CGGGGAAAAATCTAGATTCGAATTTGGCCAGACTTCCAGTTGGTGTAGAGAGATTGCG 60
1 CGGGGAAAAATCTAGATTCGAATTTGGCCAGACTTCCAGTTGGTGTAGAGAGATTGCG 60
61 GACTGCCAATGACAGTGGCTGAGTAAAGAAACAGCTTACTGATGAGCAATTACTAGAACAT 120
61 GACTGCCAATGACAGTGGCTGAGTAAAGAAACAGCTTACTGATGAGCAATTACTAGAACAT 120
121 TTCAAAAGTCTGGGATTAAGGCGCAAGTCCACTTTGAAATTAAGCTTTAGGCTCGTGTG 180
121 TTCAAAAGTCTGGGATTAAGGCGCAAGTCCACTTTGAAATTAAGCTTTAGGCTCGTGTG 180
181 GGGGCTTTTGTGTTTATTTATTTCAACTGCAAGTGGTCCATCCCTATATCAATTTAA 240
181 GGGGCTTTTGTGTTTATTTATTTCAACTGCAAGTGGTCCATCCCTATATCAATTTAA 240
241 GACGAAATTTATATCAATTCATGCGATCCCATCCCAATTAACCTGCTCTATATTAA 300
241 GACGAAATTTATATCAATTCATGCGATCCCATCCCAATTAACCTGCTCTATATTAA 300
301 TTAATTAAGAACTATCAAGAGCTTCTTATCAATTTATACATATCCAAAGAGGAGACG 360
301 TTAATTAAGAACTATCAAGAGCTTCTTATCAATTTATACATATCCAAAGAGGAGACG 360
361 ATGAACCTTCAAAAGATTGGCGAAAAAAGCAGCTGCCATTCAGAGTGTATATTA 420
361 ATGAACCTTCAAAAGATTGGCGAAAAAAGCAGCTGCCATTCAGAGTGTATATTA 420
421 GTAGAGGAGGAGCGGCGCATATTTTTCGAGCAAAATGAACAGCGGGGACTTAAGAA 480
421 GTAGAGGAGGAGCGGCGCATATTTTTCGAGCAAAATGAACAGCGGGGACTTAAGAA 480
481 GACTATGCTTTTGGCCATATTACAGCGCTGACATGCTAAATAATTCAGAGCAACAAAC 540
481 GACTATGCTTTTGGCCATATTACAGCGCTGACATGCTAAATAATTCAGAGCAACAAAC 540
541 AGTCCTCAATTTAAAGTCCCTCAATTCATGATGATGATGATGATGATGATGATGATG 600
541 AGTCCTCAATTTAAAGTCCCTCAATTCATGATGATGATGATGATGATGATGATGATG 600
601 AAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
601 AAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
661 CAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
661 CAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
721 GGTGACCCCAAAACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 780
721 GGTGACCCCAAAACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 780
781 ACATCGATTGACAGCTGAAAAAATGCTGGAAGAGATTTGAAGATATGAGATTAATTTG 840
781 ACATCGATTGACAGCTGAAAAAATGCTGGAAGAGATTTGAAGATATGAGATTAATTTG 840
841 CCAATGATCGCTATCTTAATATCAACACAGAGAGTGTAGTGTCTGCTACTTTAAC 900
841 CCAATGATCGCTATCTTAATATCAACACAGAGAGTGTAGTGTCTGCTACTTTAAC 900
901 AAAGATGCGCAAGTCCGTTTATCTATACAGATTACTGAGTAATCTGAAGATGTTGA 960
901 AAAGATGCGCAAGTCCGTTTATCTATACAGATTACTGAGTAATCTGAAGATGTTGA 960
961 ACCGCTGCTGATCAACCAATATTTCAACCTGCTCAAGTAACCTATATCCACCGGATGCA 1020
961 ACCGCTGCTGATCAACCAATATTTCAACCTGCTCAAGTAACCTATATCCACCGGATGCA 1020

```

OY 1021 GCTACACTTAAGGATGATGATATCATCAATAATCTCTTTGATGGGAGACGGT 1080
Db 1021 GCTACACTTAAGGATGATGATATCATCAATAATCTCTTTGATGGGAGACGGT 1080
OY 1081 ACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGGCAAGTGGATTTACAGTGATTAAC 1140
Db 1081 ACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGGCAAGTGGATTTACAGTGATTAAC 1140
OY 1141 CATACTTTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 1200
Db 1141 CATACTTTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 1200
OY 1201 GCGAATTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 1260
Db 1201 GCGAATTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 1260
OY 1261 TATGGCGGAGTGAAGTCTCTTCCAGATGAAGAAAAATAAATCTCTCAAAAGTCTTAAA 1320
Db 1261 TATGGCGGAGTGAAGTCTCTTCCAGATGAAGAAAAATAAATCTCTCAAAAGTCTTAAA 1320
OY 1321 AAACAATTTGCTCTTTAGCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 AAACAATTTGCTCTTTAGCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
OY 1381 ACAGTGAAGAGTGTATGAACCATTAAGTGCATCAACACAGTACAGATGAAGTGAAC 1440
Db 1381 ACAGTGAAGAGTGTATGAACCATTAAGTGCATCAACACAGTACAGATGAAGTGAAC 1440
OY 1441 CGCGCAATATATTTAAATGAATTAATGAATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 CGCGCAATATATTTAAATGAATTAATGAATGATGATGATGATGATGATGATGATGATGAT 1500
OY 1501 AAAATGAGAGATGATGATTAAGCAAAAGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 AAAATGAGAGATGATGATTAAGCAAAAGATGATGATGATGATGATGATGATGATGATGAT 1560
OY 1561 TCCTTAATATGCGCACACACACCCGATTAATGAATGAATGATGATGATGATGATGATGAT 1620
Db 1561 TCCTTAATATGCGCACACACACCCGATTAATGAATGAATGATGATGATGATGATGATGAT 1620
OY 1621 GACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 GACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
OY 1681 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
OY 1741 CTGCGGAGACAGCTTGGGGTTAATTAAGGGTCTGACACATCTGGAGAGAAAAATAGT 1800
Db 1741 CTGCGGAGACAGCTTGGGGTTAATTAAGGGTCTGACACATCTGGAGAGAAAAATAGT 1800
OY 1801 TCCGGAACAAGACATTCCTCATAGGATATCTCCCAATTAAGAAATGTCACGTGCA 1860
Db 1801 TCCGGAACAAGACATTCCTCATAGGATATCTCCCAATTAAGAAATGTCACGTGCA 1860
OY 1861 GATCTACTCTTCCCAACCTTTTTTTTAAATAAAGGTTTGTACACCTTTAGAA 1920
Db 1861 GATCTACTCTTCCCAACCTTTTTTTTAAATAAAGGTTTGTACACCTTTAGAA 1920
OY 1921 GAAAAAAGACATCTTGGCCCGGCAATTAACCATTTCCGAAGAGCTACCAATTA 1980
Db 1921 GAAAAAAGACATCTTGGCCCGGCAATTAACCATTTCCGAAGAGCTACCAATTA 1980
OY 1981 AGAATTTTCTTTTATGAGGGGAGTCTTTTCTATCTTGGGGTGGGATTTG 2040
Db 1981 AGAATTTTCTTTTATGAGGGGAGTCTTTTCTATCTTGGGGTGGGATTTG 2040
OY 2041 GCGCCACAGGAGGACCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
Db 2041 GCGCCACAGGAGGACCTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
OY 2101 CTATATATCAAGCGCCGCAAAATTAATTAACCAAGAAATCTTAAGATGAGCGGT 2160

```

```

Db 2101 CTATATATCAAGCGCCGCAAAATTAATTAACCAAGAAATCTTAAGATGAGCGGT 2160
OY 2161 ACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGGCAAGTGGATTTACAGTGATTAAC 2220
Db 2161 ACAGTTTATCAAAATATTCAGCAATTTATCGATGAAGGCAAGTGGATTTACAGTGATTAAC 2220
OY 2221 CCGAATTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 2280
Db 2221 CCGAATTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 2280
OY 2281 TTACCTTTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 2340
Db 2281 TTACCTTTAAGAGACCCCTCAGTATGTGAAGATTAAGGCCCAATAATATCTGTCTTTGAA 2340
OY 2341 CGAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db 2341 CGAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
OY 2401 TGGTGATC 2408
Db 2401 TGGTGATC 2408

RESULT 3
AR208488
LOCUS AR208488
DEFINITION Sequence 5 from patent US 6383769.
ACCESSION AR208488
VERSION AR208488.1 GI:21509655
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2408)
AUTHORS Tsusaki,K., Kubota,M., and Chaen,H.
TITLE Polypeptides having .beta.-fructofuranosidase activity
JOURNAL Patent: US 6383769-A 5 07-MAY-2002;
FEATURES
source 1..2408
BASE COUNT 777 a 492 c 496 g 643 t
ORIGIN

Query Match 100.0%; Score 2408; DB 6; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGGGGAAAAATCTAGATTTCCAAATTTGGCCAGACTTCCAGTTGGTGTAGAGAGATTGG 60
Db 1 CGGGGAAAAATCTAGATTTCCAAATTTGGCCAGACTTCCAGTTGGTGTAGAGAGATTGG 60
OY 61 GACTGCCAATGACAGTGTGGCGTAAGAAAACAGCTTACTCATAGAGCAATTAATTAATTAAT 120
Db 61 GACTGCCAATGACAGTGTGGCGTAAGAAAACAGCTTACTCATAGAGCAATTAATTAATTAAT 120
OY 121 TTCAAAAGTCTGGGATTAAGGCCAAGTCCACTTTGAATTAATTAATTAATTAATTAATTA 180
Db 121 TTCAAAAGTCTGGGATTAAGGCCAAGTCCACTTTGAATTAATTAATTAATTAATTAATTA 180
OY 181 GGGGCTTTTGTGTTTATTTATTTCAACTGCAAGTGTCCATCCCTATATCAATTTAA 240
Db 181 GGGGCTTTTGTGTTTATTTATTTCAACTGCAAGTGTCCATCCCTATATCAATTTAA 240
OY 241 GACGAATTTCAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 GACGAATTTCAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
OY 301 TTAATTAAGAACTATCAAGAGCTTTCTATCAATTAATTAATTAATTAATTAATTAATTAAT 360
Db 301 TTAATTAAGAACTATCAAGAGCTTTCTATCAATTAATTAATTAATTAATTAATTAATTAAT 360
OY 361 ATGAATTTCAAAAGATTTGGCGAAAAAAGACAGCTGCGGTAACTTCAGAGACTGTATATTA 420

```


KEYWORDS JP 199806586-A/2.
SOURCE Bacillus sp.
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 2408)
AUTHORS Tsusaki, K., Kubota, M. and Chaen, H.
TITLE POLYPEPTIDE HAVING BETA-FRUCTOFURANOSIDASE ACTIVITY
JOURNAL Patent: JP 199806586-A 2 10-MAR-1998;
HAYASHIBARA BIOCHEM LAB INC
COMMENT OS Bacillus sp.
PN JP 199806586-A/2
PD 10-MAR-1998
PE 09-JUN-1997 JP 1997164875
PR 10-JUN-1996 JP 96P 170630
PI TSUSAKI KEIJI, KUBOTA MICHIO, CHAEN HIROTO
PC C12N15/09, C07H21/04, C07K14/32, C12N1/21, C12N9/10, C12N9/24, PC
(C12N15/09, C12N1/07), (C12N1/21, C12R1/19), (C12N9/24, C12R1/19); CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..2408
FT /organism='Bacillus sp.'
FT /strain='V230'
FT CDS 361..1824
FT /product='beta-fructofuranosidase' FT
FEATURES
source location/Qualifiers
1..2408
/organism='Bacillus sp.'
/db_xref='taxon:1409'
BASE COUNT 777 a 492 c 496 g 643 t
ORIGIN
Query Match 100.0%; Score 2408; DB 6; Length 2408;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGGGAAATACAGATGATTCGAATGGCCAGACCTCCAGTGGGTAGAGAGAGTTCG 60
DB 1 CGGGGAAATACAGATGATTCGAATGGCCAGACCTCCAGTGGGTAGAGAGAGTTCG 60
QY 61 GACTGCCAATGCAGCTGTGGCTAGAGAAACAGCTTACTCATGAGCAATTAAGAGAAAT 120
DB 61 GACTGCCAATGCAGCTGTGGCTAGAGAAACAGCTTACTCATGAGCAATTAAGAGAAAT 120
QY 121 TTCAAAAGTCTGGGATTAAGGCCAAGTCCACTTTGAATTAACCTTTTCACCTCTGTGTG 180
DB 121 TTCAAAAGTCTGGGATTAAGGCCAAGTCCACTTTGAATTAACCTTTTCACCTCTGTGTG 180
QY 181 GGGGCTTTTGTATTTATTTATTTCAATGCAAGTGTCCATCCCTATATCAATTTAA 240
DB 181 GGGGCTTTTGTATTTATTTATTTCAATGCAAGTGTCCATCCCTATATCAATTTAA 240
QY 241 GAGCAATTTCAATCAATCCATGCCATCCCAATTAACCTGCTCCTCTATACTTTAA 300
DB 241 GAGCAATTTCAATCAATCCATGCCATCCCAATTAACCTGCTCCTCTATACTTTAA 300
QY 301 TTATATAGAAACATCAAGAGCTTTCTTATCAATTTATATCAATTCACAGAGAGAGC 360
DB 301 TTATATAGAAACATCAAGAGCTTTCTTATCAATTTATATCAATTCACAGAGAGAGC 360
QY 361 ATGAACCTCAAAAGATTGGGAAAAAAGAGCTGGCTAGCACTTCAGAGACTATTTTA 420
DB 361 ATGAACCTCAAAAGATTGGGAAAAAAGAGCTGGCTAGCACTTCAGAGACTATTTTA 420
QY 421 GTAGAGCGGAGCGAGCGCATATTTTTCGACAGCAATTAAGAGCGGAGCTACAAGAA 480
DB 421 GTAGAGCGGAGCGAGCGCATATTTTTCGACAGCAATTAAGAGCGGAGCTACAAGAA 480
QY 481 GACTATGTTTGGCCATTAATACAGCGGCTGACATGCTAAATAATCCAGAGACAACAAAC 540
DB 481 GACTATGTTTGGCCATTAATTAACAGCGGCTGACATGCTAAATAATCCAGAGACAACAAAC 540

DB 481 GACTATGTTTGGCCATTAATTAACAGCGGCTGACATGCTAAATAATCCAGAGACAACAAAC 540
QY 541 AGTCCCAATTTAAAGTCCCTCAATTCATAGCATTCAGCAATCAAAAACATTCGATTCGCA 600
DB 541 AGTCCCAATTTAAAGTCCCTCAATTCATAGCATTCAGCAATCAAAAACATTCGATTCGCA 600
QY 601 AAAGGATATATAGTACAGCAACTTAATAGATTTAGATTAAGGATAGCTGGCCACTG 660
DB 601 AAAGGATATATAGTACAGCAACTTAATAGATTTAGATTAAGGATAGCTGGCCACTG 660
QY 661 CAAAAGCGTATGTGATCGCGCAATTTATCATGATATCATCTCTCCCTTTAGCA 720
DB 661 CAAAAGCGTATGTGATCGCGCAATTTATCATGATATCATCTCTCCCTTTAGCA 720
QY 721 GGTGACCCAAAACAGTGTATGATCCCTCATTTATTCATCAAAAAGTGGTGTAT 780
DB 721 GGTGACCCAAAACAGTGTATGATCCCTCATTTATTCATCAAAAAGTGGTGTAT 780
QY 781 ACATCGATTCACAGCTGGAAAAAATGCTGAAGAGATTTGAAAGATATGATTAATTTGTT 840
DB 781 ACATCGATTCACAGCTGGAAAAAATGCTGAAGAGATTTGAAAGATATGATTAATTTGTT 840
QY 841 CCAATATGATCCGATATCTTAATATCAAAACACAGAGTGTGACGTTCTGCTTTAAC 900
DB 841 CCAATATGATCCGATATCTTAATATCAAAACACAGAGTGTGACGTTCTGCTTTAAC 900
QY 901 AAAGATGGCCAGAGCTGTTATTTCTATACAGATTACTCAGGTAACTCTGAAGTGGTGA 960
DB 901 AAAGATGGCCAGAGCTGTTATTTCTATACAGATTACTCAGGTAACTCTGAAGTGGTGA 960
QY 961 ACCGCTGCTGTAAACCAATCATTTCAACGCTCAAGTAACATTTCCAGCCGGATGCA 1020
DB 961 ACCGCTGCTGTAAACCAATCATTTCAACGCTCAAGTAACATTTCCAGCCGGATGCA 1020
QY 1021 GCTACACTTAAAGTCCATGATGATGATCATTAATCTGCTTTGATGGCGAGACGCT 1080
DB 1021 GCTACACTTAAAGTCCATGATGATGATGATCATTAATCTGCTTTGATGGCGAGACGCT 1080
QY 1081 ACAGTTTATCAAAATTTTCGCAATTTATGATGATGAAGGCAAGTGGATTCAGGTGAATAC 1140
DB 1081 ACAGTTTATCAAAATTTTCGCAATTTATGATGATGAAGGCAAGTGGATTCAGGTGAATAC 1140
QY 1141 CATACTTTAAGAGACCTCCTACTATGTTGAAGATTAAGGCCATTAATATCTTCTTTGAA 1200
DB 1141 CATACTTTAAGAGACCTCCTACTATGTTGAAGATTAAGGCCATTAATATCTTCTTTGAA 1200
QY 1201 GCGAATATCTGGAACACAGATGTTATCAAGGGGATTCAGCTTTCATATAAAGCTTAC 1260
DB 1201 GCGAATATCTGGAACACAGATGTTATCAAGGGGATTCAGCTTTCATATAAAGCTTAC 1260
QY 1261 TATGGCGAAGTGAAGCTCTCTCCAGAAATGAAGAAAAATTAACGCTTCAAACTCTAA 1320
DB 1261 TATGGCGAAGTGAAGCTCTCTCTCCAGAAATGAAGAAAAATTAACGCTTCAAACTCTAA 1320
QY 1321 AAACAAATTTCTTTAGCGAATGTTGCAATTAAGGCTGTTGAATTTGGCCGATGACTAT 1380
DB 1321 AAACAAATTTCTTTAGCGAATGTTGCAATTAAGGCTGTTGAATTTGGCCGATGACTAT 1380
QY 1381 ACAGTGAAGAGTGTATGAAGACATTAAGTGCATCAAAACCAATACAGATGAAGTGCAG 1440
DB 1381 ACAGTGAAGAGTGTATGAAGACATTAAGTGCATCAAAACCAATACAGATGAAGTGCAG 1440
QY 1441 CGGCGCAATTAATTTAAATGAATTAATGATATCTATTCACGGAATTCAGAGATTC 1500
DB 1441 CGGCGCAATTAATTTAAATGAATTAATGATATCTATTCACGGAATTCAGAGATTC 1500
QY 1501 AAAATGACAGTATGATTAAGCAAAAGATGTTTATATGCTAGAGCCCGGAGCGAC 1560
DB 1501 AAAATGACAGTATGATTAAGCAAAAGATGTTTATATGCTAGAGCCCGGAGCGAC 1560
QY 1561 TCCTTAATATGCGCCACACAAACCCGATTAATGAAGTGAAGTGAAGTGAAGTGAAGT 1620
DB 1561 TCCTTAATATGCGCCACACAAACCCGATTAATGAAGTGAAGTGAAGTGAAGTGAAGT 1620

1621 GACCTGCTGATCTACACACACTTACTTCATTCGCGTATCCCGACCCCTGAAGTAAAT 1680
1621 GACCTGCTGATCTACACACACTTACTTCATTCGCGTATCCCGACCCCTGAAGTAAAT 1680
1681 AATGTGACTCTACACACTTATATAGACGATAGAGGCTTCTATCCGAACATCACTCTCAC 1740
1681 AATGTGACTCTACACACTTATATAGACGATAGAGGCTTCTATCCGAACATCACTCTCAC 1740
1741 CTGGGGGACAGCTGGGGCTAATATTAAGGCTGTGACATCTGGAGCGAATAATAGT 1800
1741 CTGGGGGACAGCTGGGGCTAATATTAAGGCTGTGACATCTGGAGCGAATAATAGT 1800
1801 TCCGGACAGAGCAATTCCTACAGATTAATCTCCCAATAAAGAAATGACACTGGCAA 1860
1801 TCCGGACAGAGCAATTCCTACAGATTAATCTCCCAATAAAGAAATGACACTGGCAA 1860
1861 GATCCTACCTTTTCCCAACCTTTTAAAAAAGGGTTTGTACACCTTTAGAA 1920
1861 GATCCTACCTTTTCCCAACCTTTTAAAAAAGGGTTTGTACACCTTTAGAA 1920
1921 GAAAAAAGAAATCTTGGCGGGGCAATTAACCAATGCGGAAGAGCTAACCCATTAATA 1980
1921 GAAAAAAGAAATCTTGGCGGGGCAATTAACCAATGCGGAAGAGCTAACCCATTAATA 1980
1981 AGAAATTTTCCCTTTTAAAGGGGAGTCTTTTCTATCTTGGGGTGGGAGATTGTG 2040
1981 AGAAATTTTCCCTTTTAAAGGGGAGTCTTTTCTATCTTGGGGTGGGAGATTGTG 2040
2041 GCGCCACACAGGAGACCTTTATTAATTAATAGCATGAATCTGGCGATTTCCTGGCCTA 2100
2041 GCGCCACACAGGAGACCTTTATTAATTAATAGCATGAATCTGGCGATTTCCTGGCCTA 2100
2101 CTATTTATCCAAAGCGCGCCCAAAATATTAACCAAGGAATCTTAAGAGTGGGGGT 2160
2101 CTATTTATCCAAAGCGCGCCCAAAATATTAACCAAGGAATCTTAAGAGTGGGGGT 2160
2161 ACAGTTTACCAACACCTGCTGGAGGGAGTATGCGCTCAATACCACTTTCTGTTC 2220
2161 ACAGTTTACCAACACCTGCTGGAGGGAGTATGCGCTCAATACCACTTTCTGTTC 2220
2221 CCGCAAAATGGAAAAATGATCCGAGAGGCCCATCTTTTGGGGGAAGTATCATTAATA 2280
2221 CCGCAAAATGGAAAAATGATCCGAGAGGCCCATCTTTTGGGGGAAGTATCATTAATA 2280
2281 TTACCTTTACCAACAAAGACTATCCAGATGAATGGTACGAATGGGGCAATGCAAGCTC 2340
2281 TTACCTTTACCAACAAAGACTATCCAGATGAATGGTACGAATGGGGCAATGCAAGCTC 2340
2341 CGAAGATTATTTGATGGAGCGAGCAAGGATTCGCCATCCGAAGTATACCAATAAATA 2400
2341 CGAAGATTATTTGATGGAGCGAGCAAGGATTCGCCATCCGAAGTATACCAATAAATA 2400
2401 TGGTGATC 2408
2401 TGGTGATC 2408

RESULT 5
LOCUS AR078181 1365 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 4 from patent US 5962297.
ACCESSION AR078181
VERSION AR078181.1 GI:10004927
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1365)
AUTHORS Tsusaki,K., Kubota,M. and Chaen,H.
TITLE Polypeptides having .beta.-fructofuranosidase activity
JOURNAL Patent: US 5962297-A 4 05-OCT-1999;
FEATURES Location/Qualifiers

source 1. 1365 /organism="unknown"
BASE COUNT 459 a 266 c 283 g 357 t
ORIGIN
Query Match 56.7%; Score 1365; DB 6; Length 1365;
Best Local Similarity 100.0%; Pred. No. 4.6e-299;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
457 ATGAACAGCGGGGACTACAGAAAGACTATGCTTTGGCCATATTACACCGCGCTGACATG 516
1 ATGAACAGCGGGGACTACAGAAAGACTATGCTTTGGCCATATTACACCGCGCTGACATG 60
517 CTAAAAATTCAGACAAACAACAAACAGTCCATTTTAAAGGCGCTCAATTCATGATCA 576
61 CTAAAAATTCAGACAAACAACAAACAGTCCATTTTAAAGGCGCTCAATTCATGATCA 120
577 GCAATCAAAAACATTGATTGGCAAAAAGGATAGTAAGTCAGGCAACCTTAATAGATT 636
121 GCAATCAAAAACATTGATTGGCAAAAAGGATAGTAAGTCAGGCAACCTTAATAGATT 180
637 GATGATAGGATAGTACCTGGCCACTGCCAAACCGCTGATGCTACCGGCAATTATCATGA 696
181 GATGATAGGATAGTACCTGGCCACTGCCAAACCGCTGATGCTACCGGCAAAATTATCATGA 240
697 TATCATCATGCTCCCGCTTTAGACAGGTGACCCAAAAAACAGTATGATACCTCAT 756
241 TATCATCATGCTCCCGCTTTAGACAGGTGACCCAAAAAACAGTATGATACCTCAT 300
757 TTATTTCTATCAAAAAGTGGTATGATCATGATTCAGACCTGAAAAATCTGGAAGAT 816
301 TTATTTCTATCAAAAAGTGGTATGATCATGATTCAGACCTGAAAAATCTGGAAGAT 360
817 TTGAAGATATGATGATTTTGTTCCTCAATATGATCCGTTTAAATATCAACACAGAG 876
361 TTGAAGATATGATGATTTTGTTCCTCAATATGATCCGTTTAAATATCAACACAGAG 420
877 TGGTCAGGTTCTGCTACTTTAACCAAGATGGCCAAAGCGCTTATCTATACAGATTAC 936
421 TGGTCAGGTTCTGCTACTTTAACCAAGATGGCCAAAGCGCTTATCTATACAGATTAC 480
937 TCAGCTAATCTGTAAGATGTGGAACCGGTCTGTGTAACCAATCATTTCACTGCTCA 996
481 TCAGCTAATCTGTAAGATGTGGAACCGGTCTGTGTAACCAATCATTTCACTGCTCA 540
997 GTAACCTATCCCAAGCGGATGACGCTAACCTTAAAGTCGATGAGATATGATCAATA 1056
541 GTAACCTATCCCAAGCGGATGACGCTAACCTTAAAGTCGATGAGATATGATCAATA 600
1057 TCTGCTTTGATGGCGAGACGCTATACGTTATCAAAATATTACGAATTTATGATGAA 1116
601 TCTGCTTTGATGGCGAGACGCTATACGTTATCAAAATATTACGAATTTATGATGAA 660
1117 GCGAAGTGATTTGAGTGATTAACCATTTTAAAGAGCCCTCACTATGTTGAAGATAAG 1176
661 GCGAAGTGATTTGAGTGATTAACCATTTTAAAGAGCCCTCACTATGTTGAAGATAAG 720
1177 GCGCAATATATCTGCTTTGAGCGGAATCTGGAACCAACAGATGTTATCAAGCGAT 1236
721 GCGCAATATATCTGCTTTGAGCGGAATCTGGAACCAACAGATGTTATCAAGCGAT 780
1237 CAGTCTTCAATTAATTAAGCTTACATGCGGAGTACGCTCTCTCCAGAAATGAATAA 1296
781 CAGTCTTCAATTAATTAAGCTTACATGCGGAGTACGCTCTCTCCAGAAATGAATAA 840
1297 AATAAAGCTTCAAAAGTCTTAATAAACAAATTCCTTTAGCGAATGGTGCAATTAGGC 1356
841 AATAAAGCTTCAAAAGTCTTAATAAACAAATTCCTTTAGCGAATGGTGCAATTAGGC 900
1357 ATTGTTGAATGGCGGATGATATACAGTGAAGGTATGAACAACATTATGTCGATCA 1416
901 ATTGTTGAATGGCGGATGATATACAGTGAAGGTATGAACAACATTATGTCGATCA 960

1417 AACACAGTAGAGATGAAGTGCAGACGCCCAATATATTTAAATGAATTAATGATAT 1476
1417 AACACAGTAGAGATGAAGTGCAGACGCCCAATATATTTAAATGAATTAATGATAT 1476
961 AACACAGTAGAGATGAAGTGCAGACGCCCAATATATTTAAATGAATTAATGATAT 1020
1477 CTATTCAGGAGTTTCAAGAGATCCAAATGAAGTATGATGAATTAACGACAAAGATGT 1536
1021 CTATTCAGGAGTTTCAAGAGATCCAAATGAAGTATGATGAATTAACGACAAAGATGT 1080
1537 TATATGAGGAGCCCGAGGAGGAGCTCTTAATGAGCCACACAAACCCGATTAATGAAC 1596
1081 TATATGAGGAGCCCGAGGAGGAGCTCTTAATGAGCCACACAAACCCGATTAATGAAC 1140
1597 GGAATGATTAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1656
1141 GGACTGATTAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
1657 GGTATCCCGACCCGAGAGGATTAATGATGATGATGATGATGATGATGATGATGATG 1716
1201 GGTATCCCGACCCGAGAGGATTAATGATGATGATGATGATGATGATGATGATGATG 1260
1717 TTCTATCCAGAACATCACTCTACCTGCGGACAGAGTTGGGTTAATTAAGGATCT 1776
1261 TTCTATCCAGAACATCACTCTACCTGCGGACAGAGTTGGGTTAATTAAGGATCT 1320
1777 GACACATCTGGAGAGAAATAGTTCCGACAGAGACAAATTC 1821
1321 GACACATCTGGAGAGAAATAGTTCCGACAGAGACAAATTC 1365

RESULT 6
AR208487
LOCUS AR208487 1365 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 4 from patent US 6383769.
ACCESSION AR208487
VERSION AR208487.1 GI:21509653
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1365)
AUTHORS Tsusaki,K., Kubota,M. and Chaen,H.
TITLE Polypeptides having .beta.-fructofuranosidase activity
JOURNAL Patent: US 6383769-A 4 07-MAY-2002;
FEATURES
source 1..1365
BASE COUNT 459 a 266 c 283 g 357 t
ORIGIN

Query Match 56.7%; Score 1365; DB 6; Length 1365;
Best Local Similarity 100.0%; Pred. No. 4.0e-299;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1457 ATGAACACGGGGGACCTACAGAGAGCTATGTTTGGCCATATATACAGCGGCTGACATG 516
1 ATGAACACGGGGGACCTACAGAGAGCTATGTTTGGCCATATATACAGCGGCTGACATG 60
517 CTAAAAATTTCCAGGACAAACAAAGCTCTCAATTTAAGTGCCTCAATTCATGATCA 576
61 CTAAAAATTTCCAGGACAAACAAAGCTCTCAATTTAAGTGCCTCAATTCATGATCA 120
577 GCAATCAAAAACATTTGATGCGCAAAAGGATGATTAAGTACAGGCACTTAATGATTTA 636
121 GCAATCAAAAACATTTGATGCGCAAAAGGATGATTAAGTACAGGCACTTAATGATTTA 180
637 GATGATGAGTACTGCGGACCTGCAAAAGCTGATGATGATGATGATGATGATGATGATG 696
181 GATGATGAGTACTGCGGACCTGCAAAAGCTGATGATGATGATGATGATGATGATGATG 240
697 TATCAGATCTGCTCGGCTTTAGCAGGTGACCAAAAACAGTATGATGATGATGATGATG 756
241 TATCAGATCTGCTCGGCTTTAGCAGGTGACCAAAAACAGTATGATGATGATGATGATG 300

757 TTATTTCAAAAAAGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATG 816
301 TTATTTCAAAAAAGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATG 360
817 TTTGAAGATATGATTAATTTGTTCCAAATATATATATATATATATATATATATATAT 876
361 TTTGAAGATATGATTAATTTGTTCCAAATATATATATATATATATATATATATATAT 420
877 TGGTCAGGTTGCTGCTTATTAACCAAGATGGCAAGTCCGTTTATCTATACAGATTA 936
421 TGGTCAGGTTGCTGCTTATTAACCAAGATGGCAAGTCCGTTTATCTATACAGATTA 480
937 TCAGTAACTCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
481 TCAGTAACTCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
997 GTAAACTTAATCCAGCCGATGACAGTACCTTTAAGTGCATGATGATGATGATGATG 1056
541 GTAAACTTAATCCAGCCGATGACAGTACCTTTAAGTGCATGATGATGATGATGATG 600
1057 TCTGCTTTGATGCGGAGAGCGGATACAGTTATCAAAATATTCAGCAATTTATGATGA 1116
601 TCTGCTTTGATGCGGAGAGCGGATACAGTTATCAAAATATTCAGCAATTTATGATGA 660
1117 GGCAGTGTATTCAGTATGATTAACCATCTTAAAGACCCCTCACTATGTTGAAGATAG 1176
661 GGCAGTGTATTCAGTATGATTAACCATCTTAAAGACCCCTCACTATGTTGAAGATAG 720
1177 GGCCATTAATATCTTTGTTGAAGCGGATATCTGAGAACACAGATGTTATCAAGGCTAT 1236
721 GGCCATTAATATCTTTGTTGAAGCGGATATCTGAGAACACAGATGTTATCAAGGCTAT 780
1237 CAGCTTTCAATTAATTAAGTCTTACTATGCGGAGTACGCTCTTCCAGATGA 1296
781 CAGCTTTCAATTAATTAAGTCTTACTATGCGGAGTACGCTCTTCCAGATGA 840
1297 AATAAAGCTCTTCAAGTCTTAAACCAAAATCTCTTAAAGGATGATGATGATGATG 1356
841 AATAAAGCTCTTCAAGTCTTAAACCAAAATCTCTTAAAGGATGATGATGATGATG 900
1357 ATTGTTGAATGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416
901 ATTGTTGAATGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
1417 AACACAGTAGAGATGAAGTGCAGACGCCCAATATATTTAAATGAATTAATGATGAT 1476
961 AACACAGTAGAGATGAAGTGCAGACGCCCAATATATTTAAATGAATTAATGATGAT 1020
1477 CTATTCAGGAGTTTCAAGAGATCCAAATGAAGTATGATGAATTAACGACAAAGATGT 1536
1021 CTATTCAGGAGTTTCAAGAGATCCAAATGAAGTATGATGAATTAACGACAAAGATGT 1080
1537 TATATGAGGAGCCCGAGGAGGAGCTCTTAATGAGCCACACAAACCCGATTAATGAAC 1596
1081 TATATGAGGAGCCCGAGGAGGAGCTCTTAATGAGCCACACAAACCCGATTAATGAAC 1140
1597 GGAATGATTAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1656
1141 GGACTGATTAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
1657 GGTATCCCGACCCGAGAGGATTAATGATGATGATGATGATGATGATGATGATGATG 1716
1201 GGTATCCCGACCCGAGAGGATTAATGATGATGATGATGATGATGATGATGATGATG 1260
1717 TTCTATCCAGAACATCACTCTACCTGCGGACAGAGTTGGGTTAATTAAGGATCT 1776
1261 TTCTATCCAGAACATCACTCTACCTGCGGACAGAGTTGGGTTAATTAAGGATCT 1320
1777 GACACATCTGGAGAGAAATAGTTCCGACAGAGACAAATTC 1821
1321 GACACATCTGGAGAGAAATAGTTCCGACAGAGACAAATTC 1365

RESULT 7
EL15382 1365 bp DNA linear PAT 28-JUL-1999
LOCUS Bacillus sp. gene for beta-fructofuranosidase.
DEFINITION El15382.1 GI:5710065
ACCESSION JP 199806586-A/1.
VERSION Bacillus sp.
KEYWORDS Bacillus sp.
SOURCE Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.
ORGANISM
REFERENCE 1 (bases 1 to 1365)
AUTHORS Tsusaki,K., Kubota,M. and Chaen,H.
TITLE POLYPEPTIDE HAVING BETA-FRUCTOFURANOSIDASE ACTIVITY
JOURNAL Patent: JP 199806586-A 1 10-MAR-1998;
HAYASHIBARA BIOCHEM LAB INC
COMMENT OS Bacillus sp.
PN JP 199806586-A/1
PD 10-MAR-1998
PE 09-JUN-1997 JP 1997164875
PF 10-JUN-1996 JP 96P 170630
PI TSUSAKI KEIJI, KUBOTA MICHIO, CHAEN HIROTO
PC C12N15/09, C07H21/04, C07K14/32, C12N1/21, C12N9/10, C12N9/24, PC
(C12N15/09', (C12N1/07', (C12N1/21, C12R1:19), (C12N9/24, C12R1:19); CC
strandedness: Double;
CC topology: linear;
FH Key Location/Qualifiers
FT source 1.1365 /organism='Bacillus sp.'
FT mat_peptide 1.1365 /strain='V230'
FT /product='beta-fructofuranosidase'.
FEATURES
source
1.1365 Location/Qualifiers
/organism='Bacillus sp.'
/db_xref='taxon:1409'
BASE COUNT 459 a 266 c 283 g 357 t
ORIGIN
Query Match 56.7%; Score 1365; DB 6; Length 1365;
Best Local Similarity 100.0%; Pred. No. 4.6e-299;
Matches 1365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 457 ATGAACACGGGGAGTACAGAGAGATGTTGGCCATATATACACGGCGTACATG 516
DB 1 ATGAACACGGGGAGTACAGAGAGATGTTGGCCATATATACACGGCGTACATG 60
OY 517 CTAAAAATTCAGAGCAACAAACAGTCTCAATTTAAAGTGCCTCAATTCAATGCATCA 576
DB 61 CTAAAAATTCAGAGCAACAAACAGTCTCAATTTAAAGTGCCTCAATTCAATGCATCA 120
OY 577 GCAATCAAAAACATTGATTTGGGCAAAAGGATGTAAGTCAAGCACTTAATAGATTGA 636
DB 121 GCAATCAAAAACATTGATTTGGGCAAAAGGATGTAAGTCAAGCACTTAATAGATTGA 180
OY 637 GATGATGAGTAGTGGCCACTGCAAAAAGCTGATGAGTGGCGCAAAATATCATGGA 696
DB 181 GATGATGAGTAGTGGCCACTGCAAAAAGCTGATGAGTGGCGCAAAATATCATGGA 240
OY 697 TATCAATCTGTCCTCGCTTTAGCAGGTGACCCCAAAAACAGTATGATACCTCCACTT 756
DB 241 TATCAATCTGTCCTCGCTTTAGCAGGTGACCCCAAAAACAGTATGATACCTCCACTT 300
OY 757 TTATTTCTTCAAAAAGTGGTGATATCATGATGACAGTGGAAAAATGCTGGAAGAGTA 816
DB 301 TTATTTCTTCAAAAAGTGGTGATATCATGATGACAGTGGAAAAATGCTGGAAGAGTA 360
OY 817 TTTGAAGATATGATTAATTTGTTTCCAAATGATCCGTATCTTAATATCAACACAGAGG 876
DB 361 TTTGAAGATATGATTAATTTGTTTCCAAATGATCCGTATCTTAATATCAACACAGAGG 420
OY 877 TGGTCAGGTTCTGCTACTTTTAAACCAAAAGATGCCAAGTCCGTTATTTCTATACAGATTAC 936

DB 421 TGGTCAGGTTCTGCTACTTTTAAACCAAAAGATGCCAAGTCCGTTATTTCTATACAGATTAC 480
OY 937 TCAGGTAATCCCTGAGAGATGCTGGAACCGGTGCTGTAACCAAAATATTTCAACCTCTGAA 996
DB 481 TCAGGTAATCCCTGAGAGATGCTGGAACCGGTGCTGTAACCAAAATATTTCAACCTCTGAA 540
OY 997 GTAACATATCCCAACCGGATGCAGTCACTTAAGTCGATGAGTATGATCATGATAA 1056
DB 541 GTAACATATCCCAACCGGATGCAGTCACTTAAGTCGATGAGTATGATCATGATAA 600
OY 1057 TCTGCTTTGATGGCGAGACGGTACAGTTATATCAAAATATTCACAAATTTATGCGATGAA 1116
DB 601 TCTGCTTTGATGGCGAGACGGTACAGTTATATCAAAATATTCACCAATTTATCATGATA 660
OY 1117 GGCAGTGGATTTGAGGATATACCATATCTTTAAAGACCCCTCACTATGTTGAGATGAG 1176
DB 661 GGCAGTGGATTTGAGGATATACCATATCTTTAAAGACCCCTCACTATGTTGAGATGAG 720
OY 1177 GGCATTAATATCTGCTTTTGAAGCAATATCTGGAACAACAGATGTTATCAAGCGAT 1236
DB 721 GGCATTAATATCTGCTTTTGAAGCAATATCTGGAACAACAGATGTTATCAAGCGAT 780
OY 1237 CAGTCTTTCAATATTAAGCTTACTATGCGGAGTACGCTCTTCCAGATGAAAAA 1296
DB 781 CAGTCTTTCAATATTAAGCTTACTATGCGGAGTACGCTCTTCCAGATGAAAAA 840
OY 1297 AATAACGCTCTCAAAAGCTCTTAAAAACAATGCTCTTTAGCGAATGCTGATTAAGC 1356
DB 841 AATAACGCTCTCAAAAGCTCTTAAAAACAATGCTCTTTAGCGAATGCTGATTAAGC 900
OY 1357 ATTGTTGAATGGCGAGTACATATACAGTGAAGATGTTATGAACCATATAGTCCATCA 1416
DB 901 ATTGTTGAATGGCGAGTACATATACAGTGAAGATGTTATGAACCATATAGTCCATCA 960
OY 1417 AACACAGTAGCAGATGAGTGCAGACCGCCCAATATATTTAAATGAATATTAATGTTAT 1476
DB 961 AACACAGTAGCAGATGAGTGCAGACCGCCCAATATATTTAAATGAATATTAATGTTAT 1020
OY 1477 CTATTCAGGATTTCAAGAGATCCAAATATGACGATGATGATTAATGACACAAATGTT 1536
DB 1021 CTATTCAGGATTTCAAGAGATCCAAATATGACGATGATGATTAATGACACAAATGTT 1080
OY 1537 TATATGCTAGGCGCGGAGCGAGCTCTTAATATGAGCCCAACACCGGTAATGAAT 1596
DB 1081 TATATGCTAGGCGCGGAGCGAGCTCTTAATATGAGCCCAACACCGGTAATGAAT 1140
OY 1597 GGACTTGAATTAACATGATCTTGAACCTGCTGATCTACACAGACTTACTCTCATTTGC 1656
DB 1141 GGACTTGAATTAACATGATCTTGAACCTGCTGATCTACACAGACTTACTCTCATTTGC 1200
OY 1657 GGTATCCCGGACCCCTGAAGGATATATGTTGTTACTCACAGATTATATGACGATGAGGC 1716
DB 1201 GGTATCCCGGACCCCTGAAGGATATATGTTGTTACTCACAGATTATATGACGATGAGGC 1260
OY 1717 TTCTATTCAGAACATCACTCTACCTGCGGGGCAAGCTTGGGGTTAATATTAAGGGTCT 1776
DB 1261 TTCTATTCAGAACATCACTCTACCTGCGGGGCAAGCTTGGGGTTAATATTAAGGGTCT 1320
OY 1777 GACACATCTGAGAGAAAAATAGTTCCGAGACAGCAATTTCCCA 1821
DB 1321 GACACATCTGAGAGAAAAATAGTTCCGAGACAGCAATTTCCCA 1365
RESULT 8
AE007686 11318 bp DNA linear BCT 27-JUL-2001
LOCUS Clostridium acetobutylicum ATCC824 section 174 of 356 of the
DEFINITION complete genome
ACCESSION AE007686 AE001437
VERSION AE007686.1 GI:15024739
KEYWORDS
SOURCE Clostridium acetobutylicum.

ORGANISM Clostridium acetobutylicum
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE 1 (bases 1 to 11318)
AUTHORS Nolling,J., Breton,G., Omeichenko,M.V., Markarova,K.S., Zeng,Q.,
Gibson,R., Lee,H.M., Dubois,J., Qiu,D., Hiltl,J., Wolf,Y.I.,
Tatusov,R.L., Sabathe,F., Doucette-Stamm,L., Soucaille,P.,
Daly,M.J., Bennett,G.N., Koonin,E.V. and Smith,D.R.
TITLE Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum
J. Bacteriol. 183 (16), 4823-4838 (2001)
MEDLINE 21359325
PUBMED 11466286
REFERENCE 2 (bases 1 to 11318)
AUTHORS Childress,D., Zeng,Q. and Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (24-Jun-2001) GTC Sequencing Center Production,
Flushing, and Bioinformatics teams, Genome Therapeutics Corp., 100
Beaver Street, Waltham, MA 02453-8443, USA
FEATURES
source location/Qualifiers
1. 11318
/organism="Clostridium acetobutylicum"
/strain="ATCC 824"
/db_xref="ATCC:824"
/db_xref="taxon:1488"
89. 790
/gene="CAC1770"
89. 790
/codon_start=1
/protein_id="AAK7935.1"
/product="RNA polymerase sigma factor, SigK-like"
/db_xref="GI:15024740"
/translation="MFMRSLEDRQAKESKEENRLOEKPRIASLLOKKGKYL
QYDDELSTGMAFEAEVSEYDNRKGFSPAKVITLRAIDYIRKQKRNVIYLO
GNYSDNDINPIMFNKAMEQYKNSINEIRLELLEKREMGIEFVNLVYASPK
QEKLRKYKNIANIVHMODILNKLIRVKRLPIKEIKENIYINRKKLERGRYITAIIV
IAEIGDYRLIREYE"
801. 1544
/gene="CAC1771"
801. 1544
/gene="CAC1771"
/codon_start=1
/tranyl_table=1
/product="Uncharacterized protein, Ykrl B.subtilis
homolog"
/protein_id="AAK79736.1"
/db_xref="GI:15024741"
/translation="MKGIVIELEKLYLILTEKGDFIEDRLNTRAKIGOEINIKSKK
ISTKALREPALATAILFIIINSLIAGMIFREYVVDINPNKKNVAMEIHNYF
GNIIKLOANKKGSIVKMKMKLKFATNVVINIRFANDKFIIDSGKMTYITIA
SSKSIINDESIDSLSEHYIKENKINARPMIYVINGNDYKSKQVCIPIDKTILINKVI
KNNSYKPYDLNKSIDELINIANDEMY"
1682. 3151
/gene="CAC1772"
/note="sacB"
1682. 3151
/gene="CAC1772"
/codon_start=1
/tranyl_table=1
/product="Levanucrase"
/protein_id="AAK79737.1"
/db_xref="GI:15024742"
/translation="MNKLIYKJLIGSMIGSITTOOTFASTNDMNYKFTYSHIT
RYNSKIPMEQNDLAKFYPOFNASTLKIYASAKGIDKGNLIDLDVDSWPLQNGDGT
VANKGHITVIALGDPKNDSTIYMYKIGENSIDSMNAGVFNKDSKYVANDP
YLYKQDEMSGATLISDQVRLFTYDFSGVAKDGDVDASNOVITTTQVNSODSNT
INIDSVSHKSVFPGNGTLYQNOQFIDEKWSGDNHTLRDPHYVDNGRKYLVE
ANTGTNDYOGDTSLINKAFYGRSOSFEKTEKDDLLIDTNKHDASLANGLGTELN
NDYTLKREKPLIASNTYDEIRANVFKMGKRWLYLTFDSGSKMTINGISKDIYML
GFSSNLTGPRPKLNTGLVLANINDPTDLFTYSHRAVPOTNGKNVYITSTYITNRGM
YSDHHSFAPSPFLNKTGKTJVSISNLSLQOGLTIDNY"

gene 3249. 4793
/gene="CAC1773"
/note="sacB"
3249. 4793
/gene="CAC1773"
/codon_start=1
/tranyl_table=1
/product="Levanase"
/protein_id="AAK79738.1"
/db_xref="GI:15024743"
/translation="MKRKKLKLITLITAFYIILVLSLYIQNNSTIKSTTKKON
REYVHFYVFNKKNQDPQRPVFEDEYHYLYYNQDYPNGNTEWRQATSKLVTKOE
GVCLIPKTYNENDIWSGSFVVAQNTAGFAGALVATYQPSASMDQEOPLWSTOR
GRTEKPSYNOPIINPCTKDEPDKIIMDEKNNKVMVLAETKIGFESYNLKNMOH
TGDEFTNNIGIIVECDIEPMQSDNGKNTIGTSANKVSGEPTNYAVWGVGKKE
IADISPKMIDYGFMDYAAVFPESDNGKNTKRYVALAMNNMRYANPTIQRNFC
MDSIVREITLSKQDNTYLSLSKPIKRIENITTSIDQFQSVYGLKHKYQAGAYOL
DTDINWDKRVNGVLRRESLDKRRHIDVITFEKYSYVNKAYVGNPDKSKRYVESRA
PFDINNKVHLRIEVDKSVSEVFIIDCKITYSNEVFPREDKGTITLRSIKKAVFKNI
TIKHID"
4860. 6146
/gene="CAC1774"
/note="sacB2"
4860. 6146
/gene="CAC1774"
/codon_start=1
/tranyl_table=1
/product="Levanucrase"
/protein_id="AAK79739.1"
/db_xref="GI:15024744"
/translation="MKTRKYKMTISLAVILAILTIFELLRHNTGYTSIWSROAO
FKCIKEMTAPININFKLTAPNLWMDTLPVKKDGLAVVNGKVFALYASRNVCA
NKRHDVAGISYECSTDEGMVYKGLAVVEDALSRQWAGSAILDEMGMOFEYTAG
RKGFAVRPEORLSTKTSINVDKGVHTTCSKHOVILPEPGVYQYOTMOAKPDIY
SFROPYFREDKTKKDYLIPEGKNGKGIETKKEPNIDKLRKHIAIPRGVENGV
GIAVAONKDLTRFLPLPLEAVGVNOOLEKPOVLMKKNYLETTSKRPYAGLNG
VDGLYRGCSNLSRNYRPLNGNCGLVITNPIDPYQYTSWLVLSGHDVLSFINEHFG
QLRYGCTFAPLQISLKGYSKIIIGRLEGVTPAH"
6560. 8305
/gene="CAC1775"
6560. 8305
/gene="CAC1775"
/codon_start=1
/tranyl_table=1
/product="Predicted membrane protein"
/protein_id="AAK79740.1"
/db_xref="GI:15024745"
/translation="MKIKGIIAISMSILLIGSISVYKASNNKRLONLYKEFOKLK
LNKDSNSLVRTDNNEKLSDDLKNNKVIIPNNTILNKLKPIIKSGDONSYSKTK
GNSVSAAGDSNNAPYAKLDPIYILNBSLKNGETTNTQIAMIISDPTDAGDALSA
VGSFPAVILGVLDONNGEVYQFNAGDFDLYQVTSKSKASNIYVGSIHVIPPEDY
QAFGTQSTSDVAVYAVDFSMQDAITLGTSGDMYMLKISDQSGNVQVAVY
DKRNVYINKPSSNAGICNYTTTITPTVLSNSTYNSRYLMIGKNKNDYEMLSQONT
ILDKRYTENAKOPIQETTPKNOEMEFYTAGSTVITLNNYPELRFQVROTEMLQYL
FTINDPNNASTHRSQYCGLANKYAEKRLNTTAAQDYLVIYSPSLIQDPIINDETH
LYVDGPVLIYDSTIYADNYITGPTTGSSIANINVTNVPPTAAVKDIYVSLNDVAR
VSDENYMKVAPGSEFMSDGOCCNIGIDVYVNGNNVQLMGTWRFVQAGVDPIS
IIPGLHFDVYEAGD"
8693. 9346
/gene="CAC1776"
8693. 9346
/gene="CAC1776"
/codon_start=1
/tranyl_table=1
/product="HAD superfamily hydrolase"
/protein_id="AAK79741.1"
/db_xref="GI:15024746"
/translation="MDRIIFDCYGLIDTKEPVDAISNIAKNNILNSKDAVNIYVY
EDRLAMGCVIYPRYKLIKOYVEYCEMELVTNFSKSYDEITSYKRLKPEVEVSELE
KIKHKGHIAAGYMMDIYPCSKLQMKRIYWRKKKGNKCHPEPEVHTLDEVINLI"
9598. 10531
/gene="CAC1777"

CDS 9598..10521
/gene="CAC1777"
/codon_start=1
/transl_table=11
/product="Nudix (MutT-like) hydrolase"
/protein_id="AAK79742.1"
/db_xref="GI:15024747"
/translation="MRDMNKKNGSLGLEKEFLDRLVPGDYERPSNTVMDLLTVDDVP
IEKDPNKALEKILIKRGDHPYMGAVAGFVNIWEGISSACVRELKEPTNEVNF
EOLKTEGDVNDPRPMRBYISVAMALADKSTPKAGDDADAKKRTYKRFISSGF
GIDKIDTYNILLTSDGETRIGLVTEKEKNGVVIKYPSFRLLEMSKEELFHDIE
EDICALERLKNIIEYPIAFSRPLKPYETLREAKVEEAILNLEKPLTRANFRRIKIM
VWELEKITSPPATCYTFNEDWHTFLDE"
10543..11193
/gene="CAC1778"
10543..11193
/gene="CAC1778"
/db_xref="GI:15024748"
/protein_id="AAK79743.1"
/db_xref="GI:15024748"
/translation="MGKFEKILNMESEIKGRFVDLNDLDDK/TAIYIDVNVGPH
GALSPRBEIVDEIVRINKETLGNKRIFFLDEHTNNSTEFKSYAHCLGSLAEELI
PELKNFALDSTNVMIPKNSVNGFHAPEKKELEENSOILENIIIGCEVDICVNF
NLKTYFNKNDKRIIIPSNVETDFGTHDGLMKIISLMEMOSNGIEIVRVL"
BASE COUNT 4466 a 1373 c 1945 g 3534 t
ORIGIN
Query Match 31.4%; Score 756.6; DB 1; Length 11318;
Best Local Similarity 73.1%; Pred. No. 3.3e-161;
Matches 985; Conservative 0; Mismatches 359; Indels 3; Gaps 1;
456 AATGAACGGGGGACTACAGAAGACATGTTTCCCATATATACAGCCGTGCAT 515
1765 AACTAATGACATTAATACAAAGAACATATGAGATTTGCAATATACCTGTTATATAT 1824
516 GCTAAAAATTCAGAGAACAAACAGTCTCAATTTAAAGTCCCTCAATTCATCATC 575
1825 GTCAAAAATACCAATGGAGCAAAATGACTTAAATTTAAAGTCCCTCAATTTAAAGT 1884
576 AGCAATCAAAAACATGATTGCGCAAAAGGGTATGATAGTCAAGCAACTTAATAGATT 635
1885 AACATTAATAAATATAGCTTACAGAAAGGGTATGATAGTCAAGCAACTTAATAGATT 1944
636 AGATGATGGAGATAGCGCCACTGCAAAAGCGTATGATGCGGCAATATATCATG 695
1945 GGATGATGGAGATAGCGCCACTGCAAAAGGGTATGATGCGGCAATATATCATG 2004
696 AATACACATGCTCGGCTTTAGCAGGTGACCCAAAAACAGATGATATCTCCACTTCA 755
2005 GTATCATATGTTTGGCTTTGGCAGGTGATCTTAAATAACAGATGATATCTTATTTA 2064
756 TTTATTTCTATCAAAAAGTCCGTATACATGATGATGACGTGAGAAAAGTCTGGAAGAT 815
2065 TATGTTTATCAAAAAGATTTGGAAGAACTCATTTGATGATGGAAGAACTCAGGTAAGT 2124
816 ATTGGAAGATGATGATTAATTTGTTCCAAATGATCCGTATCTTAATAAATCAACAGCA 875
2125 GTTTAAAGATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2184
876 GTGTCAGGTTCTGCTACTTTAACAAGATGCGCAAGTCTGTTATTTATATACAGTTA 935
2185 ATGCTCGTGTTCGCAACTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 2244
936 CTGAGGTAATCTGAAGATGTTGAACCGGCTGCTGATACCAATCAATTTCACTGCTCA 995
2245 TTGCGGAGTAGAAGATGTTGGAACAGATGATGATGATGATGATGATGATGATGATGAT 2304
996 AGTAACCTTATCCAGCGGATGAGTACACTTAAGTAAAGTGAAGTATGATGATGATGAT 1055
2305 AGTAACCTTATCCAGCGGATGAGTACACTTAAGTAAAGTGAAGTATGATGATGATGAT 2364

QY 1056 ATCTGCTTTGATGAGCGGAGACGCTACAGTTATTAACAATATTCAGCAATTTATCGATGA 1115
DB 2365 ATCTGCTTTGATGAGCGGAGACGCTACAGTTATTAACAATATTCAGCAATTTATCGATGA 2424
QY 1116 AGCAAGTGGATTTCAAGGTATTAACCATTTAAGAGACCCCTACTAGTTGAAGATTA 1175
DB 2425 AGGTAATGAGATGTTCTGGCATTAATCATATCTTGAGAGATCCCTCATTAATGCGAAGATTA 2484
QY 1176 GGGCCAAATATATCTGCTCTTTAAGAGATTAAGTACGACAGACAGATGTTATCAAGCGCA 1235
DB 2485 TGCGCGCAAAATATCTGTTTTCAGAGCTAACACAGAGAAACAGATGTTTTCAGAGAGCA 2544
QY 1236 TCAGCTCTTTCAATTAATTAAGCTTACTATGCGGAGAGTACGCTTCTTCAGATGAATA 1295
DB 2545 TACTTCTCTACTCATTAACCTTTTATGTTAGAGTCAAGATTTCTTTTAAACAGAGAA 2604
QY 1296 AATTA--ACTGCTTCAAGTCCCTTAAACAAATTTCTTTTAAAGGATGCTCATT 1352
DB 2605 AGATCAATTTATTAATGATGATTAATAAAGCATGATGATGATGATGATGATGATGATGAT 2664
QY 1353 AGGATGTTGATTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1412
DB 2665 AGGATGTTGATTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2724
QY 1413 ATCAACACAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1472
DB 2725 ATCAACACAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2784
QY 1473 GTATCTATTCAGGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTTCA 1532
DB 2785 GTATCTATTCAGGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTTCA 2844
QY 1533 TGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1592
DB 2845 TATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2904
QY 1593 AACTGACATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1652
DB 2905 AACTGACATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2964
QY 1653 TTGCGGATTCGCGACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1712
DB 2965 TTGCGGATTCGCGACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3024
QY 1713 AGGCTCTATTCAGGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTT 1772
DB 3025 AGGCTCTATTCAGGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTTCAAGAGATTT 3084
QY 1773 GTCTGACACATCTGAGAGAGAAATAG 1799
DB 3085 AACCAAAACATCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3111
RESULT 9
AX434229 1449 bp DNA linear PAT 28-JUN-2002
LOCUS AX434229 Sequence 2644 from Patent WO0229113.
DEFINITION AX434229
ACCESSION AX434229
VERSION AX434229.1 GI:21659037
KEYWORDS
SOURCE Bacillus licheniformis.
ORGANISM Bacillus licheniformis.
REFERENCE 1. Berka, R. and Clausen, I. G.
METHODS for monitoring multiple gene expression
PATENT: WO 0229113-A 2644 11-Apr-2002;
JOURNAL Novozymes Biotech, Inc. (US) : Novozymes A/S (DK)
FEATURES
source location/Qualifiers
1.1449
/organism="Bacillus licheniformis"
/db_xref="taxon:1402"
BASE COUNT 492 a 341 c 313 g 303 t

ORIGIN

26.0%; Score 627; DB 6; Length 1449;

Query Match Best Local Similarity 66.5%; Pred. No. 1,1e-131; Matches 957; Conservative 0; Mismatches 455; Indels 27; Gaps 3;

OY 361 ATGAATCTCAAAAGATTGGCGAAAAAGACGCTGCCGTAACCTTCAGGACTGCTATATTA 420
 DB 1 ATGAACATCAAAAACATTCCTAAAAAGCGTCACGCTTAACCGT---TGCTGGCGGCACTG 57
 OY 421 GTAGAGCGGAGCGGCATATTTTGGCCAGCAATGAACACGGGGGACTACAAAGAA 480
 DB 58 CTGGCCGAGGTGGCGCAAACTTTCAGAAAAGAGCA-----GATTAACAAA 111
 OY 481 GACTATGTTTGGCCATATTACACGGCGTGACATGCTAAATAATTCAGGACAAACAAC 540
 DB 112 AGCTAGGAGATTTCTCATATCAGACAGACATGACATGCTGAATAATTCGCGAGCAGCAAG 171
 OY 541 AGTCCTCAATTTAAAGTGCCTCAATTGATGATCAGCAATCAAAAACATTTGATTCGGA 600
 DB 172 AGCGAACAATTTAAAGTTCCTCAATTCGATCCGAAAACAATCAAAAACATCCCTTCGCA 231
 OY 601 AAGGATATAGTAGTACGCAACTTAATAGATTATGATGATGATGATGATGATGATGAT 660
 DB 232 AAGGGATATCAAAAATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 291
 OY 661 CAAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 292 CAAAATGCGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351
 OY 721 GGTGACCCAAAAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 352 GGGGATCGGAAGCGTGCATGACATCATCATCATCATCATCATCATCATCATCATCATCAT 411
 OY 781 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 412 ACTTCTATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 471
 OY 841 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 472 CCAGAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531
 OY 901 AAGATGCGCAAGTCCCTTATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 532 AAGAGCGGAAAGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 581
 OY 961 ACCGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 582 -----CGGCAAGCAGAGCGCTGACAAAGCTCAGTCAATTTCTTCAGCCGGATTGC 633
 OY 1021 GCTACACTTAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 634 GACAGCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 693
 OY 1081 ACAGTTATCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 694 ACCGTTATCAAAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
 OY 1141 CATACCTTAAGAGACCTCACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 754 CATACGATGAGAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813
 OY 1201 GCGAATACGTAACAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 DB 814 GCCAATACGGAACAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
 OY 1261 TATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 DB 874 TACGGGGGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
 OY 1321 AAACAATTCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 DB 934 AAAAGAGACGCTTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993

OY 1381 ACAGTGAAGATGTTATGAACCATTAAGTGCATCAACACAGATGATGATGATGATGATGAT 1440
 DB 994 ACAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053
 OY 1441 CCGGCCAATATATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1500
 DB 1054 CCGGCCAATCTTCAAAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1113
 OY 1501 AAATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 DB 1114 AAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
 OY 1561 TCCTTAATGCGCCACACACACACACACACACACACACACACACACACACACACACACACAC 1620
 DB 1174 TCATTAACCGGACCATTCACACACACACACACACACACACACACACACACACACACACACAC 1233
 OY 1621 GACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 DB 1234 GATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293
 OY 1681 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 DB 1294 GAAGTCTCATTAACAGCTATCATCAACACAGAGATTTGACAGCATGATGATGATGATGAT 1353
 OY 1741 CTGCGGAGCAAGCTTGGGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1799
 DB 1354 TTGACACCAAGCTTTTTCGTAAGATCAAAAGATCAAAACATCCGTTGCAAAAACAG 1412

RESULT 10
 BSSACBG
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 Bacillus subtilis.
 Reference
 1 (bases 1 to 2007)
 Authors
 Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy, J., Trepo, J. G. and Gay, P.
 The DNA sequence of the gene for the secreted Bacillus subtilis
 enzyme levansucrase and its genetic control sites
 Mol. Gen. Genet. 200 (2), 220-228 (1985)
 JOURNAL
 MEDLINE
 PUBMED
 85295507
 2993818
 REFERENCE
 PUBMED
 2 (bases 444 to 648)
 Authors
 Rouet, A., Arnaud, M., Klier, A. and Rapoport, G.
 Characterization of the precursor form of the exocellular
 levansucrase from Bacillus subtilis
 Biochem. Biophys. Res. Commun. 119 (2), 795-800 (1984)
 JOURNAL
 MEDLINE
 PUBMED
 6424671
 COMMENT
 FEATURES
 source
 Data kindly reviewed (17-Apr-1986) by M. Steinmetz.
 Location/Qualifiers
 1..2007
 /organism="Bacillus subtilis"
 /strain="Marburg"
 /db_xref="taxon:1423"
 230..235
 /note="pot. -35 region (sacB)"
 253..258
 /note="pot. -10 region (sacB)"
 294..303
 /note="put. rRNA binding site (URF)"
 305..370
 /note="stem-loop structure"
 305..332
 /note="imp. inverted repeat A"
 315..365
 /note="unidentified reading frame (aa 1-16)"
 /codon_start=1

[illegible]

QY	838	GTTCGAATGATCCGTAATCTTAATATATAACACAGGAGTGCTACGGTTCTGTACTTTA	897
Db	911	GATGCAATATGATTTATCTTAAAGACCAACACAAAGAAATGTCAGGTTTCCAGCCATTT	970
QY	898	ACCAAGATAGGCGCAGTCCGTTTATTTGATATACAGATTTCTCAGGTAATCCGAAGATGT	957
Db	971	ACATCTGACGGAAAAATCCGTTTATTTCTACCTGATTTCTCCGTAACATTA-----	1023
QY	958	GGAAACGGGTGCTGTAAACCAATCTTCAACTGCTCAAGTAACCTATPCCAGCCGGAT	1017
Db	1024	-----CGGCAAAACAAACTGAACTGCAACACTGCAAAAGTTAAACGATACGATCAGAC	1072
QY	1018	GCAGCTACACTTAAAGTGCATGAGGATATCTGATCATTAATCTGTCTTTGATGCGGAGAC	1077
Db	1073	AGCTCT---TTGAACTATACCGTGTAGAGATTTAATATCAATCTTTTGA---CGGTGAC	1126
QY	1078	GGTACAGTTTATCAAAATATTCAGCAATTTATTCGATGAAAGCCAACTGATTTACAGTAT	1137
Db	1127	GGAAAAAGTATCAAAATGTACAGAGTTCACTGATGAAAGGCAACTACAGCTCAGCCAC	1186
QY	1138	AACCTACTTTAAGAGCCCTCCTATGTTGAGATTAAGGGCATTAATATCTTCTTTT	1197
Db	1187	AACCTATAGCTGAGAGATCTCTACTAGTAGAAGATTAAGGCCAAATACTTACTATTT	1246
QY	1198	GAGCGAATACTGGAACACAGATGTTATCAAGCGCATCACTCTTTCATTAATAAGCT	1257
Db	1247	GAGCAAAACACGTGAAGATGAAGATGCTTACCAAGCCGAAGATTTTATTTAAACAGCA	1306
QY	1258	TACTATGGCGGAAGTACGCTCTTCTCCAGATGAAAAAATAACCTGCTTCAAGTCT	1317
Db	1307	TACTATGGCAAAAGACATCATCTTCCGTCAAGAACTCAAAAATCTCTCGCAAGCGAT	1366
QY	1318	AAAAACAATTTGCTCTTTACGAATGCTGATTAAGCATTTGGAATTTGGCCAGTAC	1377
Db	1367	AAAAAACCGCGCGCTGAGTTAGCAACAGCGCTCTCGGTATGATTGAGCTAAACGATAT	1426
QY	1378	TATACAGTGAANAAGTGTATGSAAACCATTTAGTCGCATCAACACAGTAGACAGTAGTTC	1437
Db	1427	TACACACTGAAAAAAGTGAAGAACCCGGATTGGCATCATAACAGTAAAGTAATTT	1486
QY	1438	GAAAGCGCCATATATTTAAATGAAATTAATATGATCTATTTACAGGATTTCAAGAGCA	1497
Db	1487	GAAAGCGGGAAGCTTTTAAATGAAAGGCAATGTTACCTGTTTACGACACTCCGCGGA	1546
QY	1498	TCCAAAATGACAGATGATGGAATTAACACAAGAAAGTTTATATGCTTAGGCGCCGAGGC	1557
Db	1547	TCAAAAATGACAGATTTGACGGCTTACGCTTAACGATATTTACATCTTGTTATGTTTTCT	1606
QY	1558	GACTCCTTAATATGCCCACACAACCCGATAAATGAACTGACTGTATTTGAACTGAAT	1617
Db	1607	AATTTCTTAACAGGCCCATATACAGCCGGGAACCAAAACTGGCCCTGTGTAAATGAT	1666
QY	1618	CTTGACCTGCTGATCTACACACACTATCTCTATTGCGGTATCCCGACCCCTGAAGT	1677
Db	1667	CTTGATCTCTAACAGATTAACCTTTTCTTACCTACACTTCGCGTACCTCAAGCGAAAGGA	1726
QY	1678	AATATATGATCTACACAAGTATATATGACGATTAAGGCTTCTTCAAGCAATCACTCT	1737
Db	1727	AACATATGCTGATTTACAAGCTATATGACAACACAGAGATTTCTACGACGACAAACAATCA	1786
QY	1738	CACCTGCGGAGACAAGCTTGGGGTTAATTTAAAGGCTGTGACACATCTGAGAGAGAAAT	1797
Db	1787	ACGTTTGGGCAAGCTTCTCTGTGAACATCAAAAGCAAGAAACATCTGTGTCAAAAC	1846
QY	1798	AGTTCCGAGACAGACAATCCCATAGGATTTATCTCCCAATAAAGAAATGTCATCG	1857
Db	1847	AG---CATCTTGAACAGACAATTAACAGTTAAACAATTAATAACGCAAAAGAAATGC	1903
QY	1858	CAAGATCTACTCTTTCCCAACCTTTTAAAAATTAAGGGTTTGTACCACTTTA	1917
Db	1904	CGATATCTCATTTGGCATTTTCTTTATTTCTTATCAACATTAAGGTGAATCCCATATGAA	1963
QY	1918	GAGGAAAAAGAA	1930

Db 811 TCAAAATGAGCATGAGCGATTACGTCTAAGCATATTACATCTGTATGTCT 752

QY 1558 GACTCCTTAATGAGCCACACACACCCGATTAATGAACTGCTGTATGACATGAT 1617

Db 751 AATCTTTTAAGTGGCCATACAGCGCTGAACAAACTGGCTGTGTAAATGAT 692

QY 1618 CTTCACCTGCTGATCAGACACACACTACTCTCATCTGATGATCCGACCCCTGAAGT 1677

Db 691 CTGATCTCTAAGCATGATACCTCTTCTACTACACTCTGCTGATCTCAGCGAAGGA 632

QY 1678 AATAATGTGTACTCACAAGTTATATGACGAATAGAGGCTTCTATCAGAACATCACTCT 1737

Db 631 AACATGTCTGATGATCAAGCATATGACAAACAGAGATCTACGACAAACATCA 572

QY 1738 CACCTGGCGGACAGCTGGGCTTATATTAAGGCTGACACATCTGACGAGAAAT 1797

Db 571 AGCTTGGCGGACAGCTTCTCTGACATCAAGGACAAACATCTGTCTCAAAAC 512

QY 1798 AGTTCGGGACAGCAATTCCTCATGAGGATATCTCCCAATTAACAAATGCTACTG 1857

Db 511 AG---CATCTTGAACAGACATTAACAGTAAACAAACGCAAAAGAAATGC 455

QY 1858 CAAGATCTTACCTTTTCCCAACCTTTTATTAATTAAGGTTTGTACCACTTTA 1917

Db 454 CGATCTCTATGCGATTTCTTTATTTCTTATCATACATAAAGGTAATCCCATATGA 395

QY 1918 GAAGAAAGAAAGA 1930

Db 394 CTATTAATAAGCA 382

RESULT 12
SYNPLA/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

SYNPLA 6347 bp DNA circular SYN 27-APR-1993
Cloning vector (pRL271) for sacB-mediated positive selection for double recombinants in gram-negative bacteria.
L05081
L05081.1 GI:209125
Cloning vector.
Cloning vector DNA.
"unidentified cloning vector
artificial sequences: vectors.
1 (bases 1 to 6347)
Cal, Y.
Molecular genetic approaches towards the understanding of heterocyst differentiation and pattern formation in the cyanobacterium Anabaena sp
unpublished (1992)
2 (sites)
Black, T.A., Cal, Y. and Wolk, C.P.
Spatial expression and autoregulation of hetR, a gene involved in the control of heterocyst development in Anabaena
unpublished (1992)
Location/Qualifiers
1..6347
/organism="unidentified cloning vector."
/db_xref="taxon:45196"
complement(473..1894)
/note="sacB gene from B. subtilis; ORF"
/codon_start=1
/transl_table=11
/protein_id="AA172302.1"
/db_xref="GI:209126"
/translation="NIKKFAKQATVLTFTTALLAGTQAFKETNPKPKYKTYGIS
HITRDMQIPEQKNEKYVPEEDSSTIKNISAKGLVWDWELQADGIVANYHG
YHIFALAGDKNADDTISIMFYKVGESIDSMNAGRVRSDKFDANDSILKDOT
QWMSGATFTDGRKILEYTFPSGKHYSKQULTTAQVAVASDSLSINTEGDEYSIF
DGDGKTQVNOOFFIDEGNYSAGDNTLDRPHVEDKGYVEFANTGTGEOYGEES
LFNKAYGKSTSPFROESOKILSDPKRTTELALANGALIMELNDYTKKWKPLIAS
NTVDEIRAVNPKMKRWLYLFTDSRSGKMTIDGTTSDYTLGLGVSNLGPYKPLN
KTGLVAKDLDPNDVTFYSHFAVQAAGNNVITVSYMTNKGFTADKOSTPAPSLN
IKGKTSVYKSDISLEQGLTVNK"
complement(2393..3127)

/note="erythromycin resistance determinant"
/codon_start=1
/transl_table=11
/protein_id="AA172303.1"
/db_xref="GI:209127"
/translation="MEKNIKHSQNFITSKHNIDKIMTINIRLNEHDNIFEIGSGH
TLELVKRCNFVTAIEIDHKLCKTTEKVDNDPOVLKDLLOEFKPNOSYKLYGNI
PYNISTDIIRKIVPDSIANEILYIEVGAFARLNTRKSLALLMAEDISILSVPR
EYFHRKPRVNSLIRLSKRSRISKDKOKTNYFPMKVNKEKKIFTKNOFNNSLKH
AGIDDLNNISEQFLSNVYLEPK"
3503..4162
/note="chloramphenicol resistance determinant"
/codon_start=1
/transl_table=11
/protein_id="AA172304.1"
/db_xref="GI:209128"
/translation="MEKRTGTYTVDISOMHREKFEAFQSOACTYNOTVOLDITAP
LKVYKRNKHFYPAFIIHLARLMAHNPFRAMKDGELIVDSVPCYTHEDQETP
SLMSEYHDDRFOLHIYSODVACYGENLAFEPKGFITNMFVSANPMVSTPDLN
ANMDFFAPVFTMGKTYTQGDVLMPLAIOVHNAVCDGFHYGRMLNELQYCDMGG
A"
4252..5221
/note="multiple cloning sites"
misc-feature
5222..6347
/note="oriV and oriT genes from plasmid pMB1"
BASE COUNT 1649 a 1305 c 1404 g 1989 t
ORIGIN
Query Match 25.4%; Score 611.8; DB 12; Length 6347;
Best Local Similarity 65.1%; Pred. No. 2.5e-128;
Matches 1024; Conservative 0; Mismatches 492; Indels 57; Gaps 6;
QY 358 ACGATGAAGCTTAAGAAATGCGGAAAGACGAGCTGCCGTACTTACGACCTCTATA 417
Db 1897 ACGATGAAGCTTAAGAAATGCGGAAAGACGAGCTGCCGTACTTACGACCTG 1838
QY 418 TTAGTAGAGGCGGAGCGACCGCATATTTTGGCGAGCAATGACGCGGAGCTACAG 477
Db 1837 CTGCGAGGAGG---CGCACTACAGCGTTTGGGAAGAAACGACAAAGCATATTAAG 1781
QY 478 GAAGACTATGTTTGGCCATATTTACAGCGCGTGAATCTAATAAATTCAGAGACAGCA 537
Db 1780 GAACATAGCGGATTTCCATATTTACAGCGCATGATGATGCTCAATCCCTGAACAGCA 1721
QY 538 AACAGTCTCAATTAAGTGGCTCAATTCATGATGATGATGATGATGATGATGATG 597
Db 1720 AAAAATGAAGAAATATCAAGTCTCAATTCATGATGATGATGATGATGATGATG 1661
QY 598 GCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 657
Db 1660 GCAAAAGG-----CTGAGAGTTTGGGACAGCTGGGCA 1628
QY 658 CTGCAAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
Db 1627 TTACAAAGGCTGAGGCGACGTGCGCAAACTATCAGCGGTACACATGCTTTCATTA 1568
QY 718 GCAGGTGACCAAAAGAGTATGATATCTCCTCATTTATTTATCAAAAGTGGCT 777
Db 1567 GCCGAGATCTCAAAAGGCGATGATGATGATGATGATGATGATGATGATGATGATG 1508
QY 778 GATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
Db 1507 GAACTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1448
QY 838 GTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 897
Db 1447 GATGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1388
QY 898 ACCAAGATGCGGACAGTCCGTTATTTATGATGATGATGATGATGATGATGATGATG 957
Db 1387 ACATGTCAGGGAAGAAATCTGTTATTTATGATGATGATGATGATGATGATGATG 1335
QY 958 GAACCGGCTGCTGATACCAATTCATTTCACTCAAGTAACTTATCCAGCGGAT 1017

Db 1720 AAAAAATGAAATATCAAGTTCCTGATTCGATTCCTCCACAATTAATAATATCTCTCT 1661
 QY 598 GCAAAAGGGATGATAGTACGACCACTTAATGATTTAGATGATAGGATACCTGGCCA 657
 Db 1660 GCAAAAGG-----CCTGGAGGTTGGGACACCTGGCCA 1628
 QY 658 CTGCAAAAGCGTATGTACTGCGGCAATTAATCATGATATCATGCTGCTCCGTTTA 717
 Db 1627 TTACAAAGAGCTGAGGCACTGTCGCAAACTATACAGGCTACACATCGCTTTGATTA 1568
 QY 718 GCAAGTGACCCCAAAAACAGTATGATCTCCACTTCATTATTCATCAAAAAGTCGT 777
 Db 1567 GCGGAGATCTCAAAATCGGATGACATGATTTACATGTTTATCAAAAAGTCGCG 1508
 QY 778 GATACATGATGACGCTGGAATAATGCTGAAGTATTTGAACAATATGATAATTT 837
 Db 1507 GAACTTTCTATGACGCTGGAATAAGCGTGGCGGCTTTTAAACAAGCGCAAAATTC 1448
 QY 838 GTTCCAAATGATCCGCTATCTTAATATCAACACAGAGTGTGAGTCTGCTACTTTA 897
 Db 1447 GATGCAAAATGATCTATCTTAAAGCAACAACAGAAATGTCAGGTTCAAGCCATTT 1388
 QY 898 ACCAAAGATGCGCAAGTCCGTTTATTTCTATACAGATTACTCAGGTAATCTGAAGTGT 957
 Db 1387 ACATCTGAGGGAATAATCCGTTTATTTCTACAGATTCTCCGTAACATTA----- 1335
 QY 958 GGAAACGGTGTGCTGTAACCAAAATCATTTCAAGTGTCAAGTAATATCCAGCGCGAT 1017
 Db 1334 -----CGGCAACAACACTGACACAGTACGCAAGTTACGATACAGCATAGAC 1286
 QY 1018 GGAGCTACACTTAAGCTGAGTATGATCATTAATCTGTTTATGAGCGGAGAC 1077
 Db 1285 AGCTCT---TTGAACATCAACGGGTGAGAGATTATAATCATCTTTCA---CGGTGAC 1232
 QY 1078 GGTACGTTTATCAAAATATTCAGCAATTTATGAGAAAGCAAGTGGATTTCAGGTGAT 1137
 Db 1231 GGAAAAACGTATCAAAATATGATGATGATGAAGGCAACTACAGCTGAGCGAC 1172
 QY 1138 AACCATCTTTTAAGACACCTGACTATGTTGAAGATGAAGGCGCATTAATATCTGCTTT 1197
 Db 1171 AACCATACCTGAGAGATCTCTCATGATGATGAAGATGAAGGCGCAAAATCTTAGATTT 1112
 QY 1198 GAAGCAATACTGGAACAACAGATGTTATCAAGGCGATCAGTCTTCAATATAAAGAT 1257
 Db 1111 GAAGCAACACTGGAACGATGAAGTGGCTACCAAGGGAAGAACTTATTTACAAAGCA 1052
 QY 1258 TACTATGGGGAAGTACGCTCTTCTCCAGATGAATAAAATAACTGCTTCAAGTCTCT 1317
 Db 1051 TACTATGGGGAAGGACATCATCTTCCGTCAGAAAGTCAAAACCTCTGCAAAACGAT 992
 QY 1318 AAAAAACAATTTGCTCTTTAGGAATGAGCATTTAGGATGTTGAATTTGGCCGATGAC 1377
 Db 991 AAAAAACGACGCTGATGATGAGCAAGGCGCTCTGCTGATGATTAAGCAATGATGAT 932
 QY 1378 TATACAGTAAAGTGTATGAACCATTAAGTCATCAACACATAGCAGATGATGATGAT 1437
 Db 931 TACACACTGAAAAAAGTGAAGACCGGATGATGATCAACACATACAGATGATGAT 872
 QY 1438 GAAGCGGCAATATATTAATTAATGAATTAATGATGATATCTATCAAGATTCAGAGGA 1497
 Db 871 GAAGCGGCAAGCTCTTTAAATGAAGCGCAATGATGATGATGATGATGATGATGAT 812
 QY 1498 TCCAAATGAGGAGTGAATTAAGCAAAAGATGATGATGATGATGATGATGATGATGAT 1557
 Db 811 TCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 752
 QY 1558 GAGCTCTTAATGAGGCGCACACACCGATTAATGAAGTGAAGTGAAGTGAAGTGAAGT 1617
 Db 751 AATTTCTTAAGTGGCCATACAAAGCGGTGAACAAACAGCGCTTGTGTTAAAAAGAT 692
 QY 1618 CTTGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1677
 Db 691 CTTGATCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632

QY 1678 AATAATGTGCTACTACAGTATATAGCAATAGAGGCTCTATCCAGACATCACTCT 1737
 Db 631 AACAAATGTCTGATTTACAACTATATGACAAACAGAGATTTACGACAGCAACATCA 572
 QY 1738 CACCTGCGGACAAAGCTTGGGGTTAATATTAAGGGTCTGACATCTGGAGAGAAAT 1797
 Db 571 ACCTTTGGCGCAAGCTTCCTCGCTGAACATCAAGGCAAGAAACATCTGTTGCAAGAC 512
 QY 1798 AGTTCCGGACAAAGACAATTTCCATAGCATTATTCGCCAATTAAGAAATGTCAC 1857
 Db 511 AG---CATCTTGAACAGAGCAATTAACAGTTAAACAAATAAAGCAAAAGAAATGC 455
 QY 1858 CAAGATGCTACCTTTTCCCAACCTTTTATTAATAAAGGTTTGTACACCTTTA 1917
 Db 454 CGATATCTTATTTGGCATTTTCTTTTATTTCTTATCAACATTAAGGTAATCCATATGA 395
 QY 1918 GAAGAAAAAAGAA 1930
 Db 394 CTATATAAAGCA 382
 RESULT 14
 AF048702/c 9297 bp DNA circular SYN 22-JUL-1998
 LOCUS AF048702
 DEFINITION Site-specific excision vector pFLP2, complete sequence.
 ACCESSION AF048702
 VERSION AF048702.1 GI:2961152
 KEYWORDS
 SOURCE
 ORGANISM
 . site-specific excision vector pFLP2.
 . site-specific excision vector pFLP2
 . artificial sequences: vectors.
 REFERENCE
 1 (bases 1 to 9297)
 Hoang,T.T., Karkhoff-Schweizer,R.R., Kutchna,A.J. and
 Schweizer,H.P.
 TITLE
 A broad-host-range FLP-FRT recombination system for site-specific
 excision of chromosomally-located DNA sequences: application for
 isolation of unmarked Pseudomonas aeruginosa mutants
 JOURNAL
 Gene 212 (1), 77-86 (1998)
 MEDLINE
 98326312
 PUBMED
 9661666
 REFERENCE
 2 (bases 1 to 9297)
 Hoang,T.T., Karkhoff-Schweizer,R.R., Kutchna,A.J. and
 Schweizer,H.P.
 AUTHORS
 Direct Submission
 Submitted (17-FEB-1998) Microbiology, Colorado State University,
 Fort Collins, CO 80523, USA
 JOURNAL
 Location/Qualifiers
 FEATURES
 source
 1. .9297
 /organism="site-specific excision vector pFLP2"
 /db_xref="taxon:73560"
 /plasmid="pFLP2"
 194. .596
 /note="region containing broad-host-range origin of
 replication from pRO1600"
 597. .1430
 /gene="rep"
 597. .1430
 /gene="rep"
 /codon_start=1
 /transl_table=11
 /product="replication protein"
 /protein_id="AAC27030.1"
 /db_xref="GI:2961153"
 /translation="MASPMVYKSNALVEAAVRLSVQEBRIYLACISQVKRSEPVTE
 VMSVTAEDIAITAGVPIESSYNQLEALRLKRRVRLTOEPNGSGKRPVITGMV
 OTIYREGGRLVETKIDMLPIETLTKOTFKYALADYAKMDSTHAIRLVELMGMD
 STIGRELETIDLRKWRQLEGRYPSIKDFRLVDPAVYVQINHSPLQVMAQRKGRK
 VTHLFSFQPKKPAKAVGKAPAKRKGKISDIEIAKQARPGETWEARARLTPMLDL
 A"
 complement(2338. .3759)
 /gene="sacB"
 complement(2338. .3759)
 gene
 CDS

Db	3762	ACGATGAACATCAAAAAGTTTGCAAAACACGAACACTATTAACTTTACTACCGCACTG	3703
Qy	418	TTAGTAGGAGCGGACGACCGCATATTTTTGGCGAGCAATGAAACAGCGGGACATACAG	477
Db	3702	CTGGGAGAGG --- CGCAACTCAAGCGTTTGCAGAAAGCAAGCAAAAGCCATATAG	3646
Qy	478	GAAGACTATGGTTTTGGCCATATTACACGGCGTACACTGCTAAATAATTCCAGGCAACAA	537
Db	3645	GAACATACGGCATTTTCCCATATTACACGGCATATATGCTGCAAAATCCCTGAACACAA	3586
Qy	538	AACAGTCCTCAATTTAAAGTGCCATCATTTCAATGACATCAGCAATCAAAAACATTTGATCG	597
Db	3585	AAAATGAAAATATATCAAGTTCCGAGTGGCATTCGTCACAAATTAATAATATCTCTCT	3526
Qy	598	GCAAAAGGATATGATAGCTACAGGCACTTAATAGATTAGATGATGGATAGCTGCCA	657
Db	3525	GCAAAAAG --- CCGTGACGTTTGGGACAGCTGGCCA	3493
Qy	658	CTGCAAAACGCTGATGGTACGCGGCAAAATATCATGGAATATCATCTGCTCCGCTTTA	717
Db	3492	TTACAAAACGCTGACGGCAGCTGTGCAAACTATACGGCTACCGCATCTTTTGCAATTA	3433
Qy	718	GCAGGTGACCCAAAAMACATGATGATACCTCCACTTATTTATCTATCAAAAAGTCGT	777
Db	3432	GCCGGAGAGTCTAAAATAATGCGGATGACACATCATGATTAACATGTTCTATCAAAAAGTGGCG	3373
Qy	778	GATCATTCGATTGACAGCTGGAAAAATGCTGGAAAGACTATTTGGAAGATATGATAAATTT	837
Db	3372	GAACCTCTATTAGACGCTGGAAAAAGCGTGGCGGCTTTTAAAGACAGCGCAAAATTC	3313
Qy	838	GTTCCAATGATCCGATATCTTAATATCAACACAGAGAGTGCTCAGGTTCTGCTACTTTA	897
Db	3312	GATGCAAAATGATTTCTATCTTAAAGACCAACACAGAAATGTCAGGTTGACGCACATTT	3253
Qy	898	ACCAAAAGATGGCCAGTCCGTTTATTCTATACAAATTACTCAGGTAATCCTGAAGATGCT	957
Db	3252	ACATCTGACGGAATAATCCGTTTATCTACCTAATTTCTCCGGTAAACATTA -----	3200
Qy	958	GGAACCGGTGCTGTGAACACAAATCATTTCACTGCTCAAGTAATATCCACCGCGAT	1017
Db	3199	-----CGGCAAAACAAACACTGACAACTGACAACTGACAAAGTATGACATCAGCTACAGC	3151
Qy	1018	GCAGCTACACTTAAAGTCGATGAGATATCTGATCATATAATCTGTTTGATGCGGAGAC	1077
Db	3150	AGCTCT --- TTGAACATCAACGGGTGAGAGATTATTAATCAAACTTTGA --- CGGTGAC	3097
Qy	1078	GGTACAGTTTTCAAAATATTCACACAATTTATCATGAAGGCAAGTGAGATTTCAGTGAT	1137
Db	3096	GGAANAAGGATCAAAATGATACACAGTCTCATCGATGAAGGCAACTACAGCTCAGCGAC	3037
Qy	1138	AACCATATCTTAAAGACCCCTCACTATGTTGAAGATTAAGGCCATTAATATCTTCTTT	1197
Db	3036	AACCATACGCTGAAGATATCTCACTAGTAGMAATTAAGGCCCAAAATCTATGATTTT	2977
Qy	1198	GAACCGAATACTGAACAACAGATGGTTATCAACAGGCACTGACTTCATTAATAAAGCT	1257
Db	2976	GAACCAAACTGGAACACTGAAGATGGCTACCAAGGCCAGAGATCTTTATTTAACAAGCA	2917
Qy	1258	TACTATGGCGGAAGTGAAGCTCTTTCAGAAATGAAAAAATAAAGCTTCAAGTCT	1317
Db	2916	TACTATGGCAAAACACATCATCTCTCCGTCAGAAAGTACAAAACCTTTCGCAAGCGAT	2857
Qy	1318	AAAAAACAAATGCTCTTTAGCAATATGATGATTAAGCATTTTGAATTTGGCGGATGAC	1377
Db	2856	AAAAAAGCGACGGGTGAGTTAGCAAAACGGCGCTTCGGTATGATTGAGCTTAAGCATGAT	2797
Qy	1378	TATACAGTGAAGAACTGTTATGAACCTATTAAGTGCATCAAAACAGACAGTGAAGATGC	1437
Db	2796	TACACACTGAAAAAAGTATGAACCCGCTGATTTGCATCTAACACATGAACAGATGAATTT	2737
Qy	1438	GAACGCGCCATATATTTAAATGAATTAATATGATATCTATTTACAGGATTAAGAGA	1497
Db	2736	GAACGCGCAACGCTTTTAAATTAACAGGCAAAATGATATCTGTTCACTGACTCCGCGGA	2677

